

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 17:08:16 ; Search time 7094.66 Seconds
(without alignments)
10695.346 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcataattttttttaaatt.....cttcctctagtgtgaatttt 2541

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	245	9.6	486	28	AQ344174
C 2	239	9.4	842	28	AQ745353
C 3	238.8	9.4	538	9	AL041450
C 4	238.6	9.4	609	9	AV760466

C	5	234.6	9.2	1030	14	CD558243
C	6	234.4	9.2	377	9	AI244254
C	7	233.2	9.2	599	9	AL702649
C	8	232.4	9.1	907	14	CF593378
C	9	232.4	9.1	1073	28	BZ610713
C	10	231.8	9.1	546	28	BZ597995
C	11	231.6	9.1	424	9	AA604843
C	12	231.6	9.1	429	28	AQ090219
C	13	231.6	9.1	703	29	AG141722
C	14	231	9.1	530	14	CA422533
C	15	230.6	9.1	547	10	AW504485
C	16	230.2	9.1	533	28	AQ037324
C	17	230	9.1	704	9	AL596482
C	18	229.8	9.0	339	10	AW270258
C	19	229.8	9.0	454	28	BZ892950
C	20	229.8	9.0	702	13	BX645944
C	21	229.6	9.0	803	28	BZ601931
C	22	229.6	9.0	2177	28	AQ839825
C	23	229	9.0	452	9	AL041894
C	24	229	9.0	583	28	AQ544177
C	25	229	9.0	1274	11	BC039325
C	26	228.8	9.0	708	28	AQ530231
C	27	228.4	9.0	759	10	AW976010
C	28	228.4	9.0	800	9	AV755512
C	29	228.2	9.0	450	28	AQ409292
C	30	228	9.0	383	10	BE139267
C	31	228	9.0	657	28	AQ782480
C	32	227.8	9.0	698	28	BZ601203
C	33	227.6	9.0	445	28	AQ349141
C	34	227.6	9.0	735	9	AL042756
C	35	227.4	8.9	487	9	AI859834
C	36	227.4	8.9	509	12	BM511372
C	37	227.4	8.9	556	12	BM510544
C	38	227.4	8.9	656	14	CB115054
C	39	227.4	8.9	771	13	BU617429
C	40	227.4	8.9	775	14	CA428724
C	41	227.2	8.9	354	28	AQ101564
C	42	227.2	8.9	465	9	AI460050
C	43	227.2	8.9	1029	28	BZ610725
C	44	227	8.9	723	13	BX642407
C	45	226.8	8.9	410	13	BX484725

ALIGNMENTS

RESULT 1
AQ344174

LOCUS

DEFINITION

RP0111-124P5.TV RPCI-11 Homo sapiens genomic clone

genomic survey sequence.

ACCESSION AQ344174

VERSION AQ344174.1

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 486)

AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL Map Building (1997)

COMMENT Unpublished (1997)

Other_GSSs: RPCI11-124P5.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

AQ344174 486 bp DNA linear GSS 07-MAY-1999
RPCI11-124P5.TV RPCI-11 Homo sapiens genomic clone RPCI-11-124P5,
genomic survey sequence.

ACCESSION AQ344174

VERSION AQ344174.1

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 486)

AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL Map Building (1997)

COMMENT Unpublished (1997)

Other_GSSs: RPCI11-124P5.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers
1..486
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7547596"
/db_xref="taxon:9606"
/clone="RPC1-11-124P5"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPC1-11"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; RPC111 Human Male BAC Library"

ORIGIN

Query Match 9.6%; Score 245; DB 28; Length 486;
Best Local Similarity 87.1%; Pred. No. 2.8e-27;
Matches 269; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 964 TTTTGTGTTGTTGACCGAGTCTTACTCTGTCAACCCAGGCTGAGTGTATGTCATG 1023
Db 8 TTCTTTTGTGTTGACCGAGTCTTACTCTGTCAACCCAGGCTGAGTGTATGTCATG 67
QY 1024 GTCTCAGTCTACTGCAACCTCCGCTCCGGGTTCAAGCATTTCTCCACCTCAGCTCC 1083
Db 68 GTCTCAGTCTACTGCAACCTCTGCTCTCCGGTTCAAGCATTTCTCCACCTCAGCTCC 127
QY 1084 CAGTAGCTGGGACTACAGCATGAGCCAGCCGCGCTAATTTTGTATTTTAGTA 1143
Db 128 CGAGTAGCTGGGACTACAGCATGAGCCAGCCGCGCTAATTTTGTATTTTAGTA 187
QY 1144 GAGACGTGGTTCCACTATGTTGCCAGGCTGATCTCGAACTCTGACCTTGTATCCACC 1203
Db 188 GAGACGGGTTTCACTATGTTGCCGCGCTGCTTGAACCTCTGACCTCGTATCCACC 247
QY 1204 CGCTCGGCTGCCAAAGTCTGGGATTACAGCGTGAACCTTGTGCTTGGCCGATTTT 1263
Db 248 TGCTCGGCTGCCAAAGTCTGGGATTACAGCGTGAACCTTGTGCTTGGCCGATTTT 307
QY 1264 TTAATAAAT 1272
Db 308 TTCTATAAT 316

RESULT 2

AQ745353/c
LOCUS
DEFINITION
HS_2276_A1_C05_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2276 Col=9 Row=E, genomic survey
ACCESSION
AQ745353
VERSION
AQ745353.1 GI:5522875
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 842)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
AUTHORS
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
PUBMED
10449764
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 2276 row: E column: 9
Seq primer: T7
Class: BAC ends
High quality sequence stop: 842.

FEATURES

Location/Qualifiers
1..842
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2276 Col=9 Row=E"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 9.4%; Score 239; DB 28; Length 842;
Best Local Similarity 83.2%; Pred. No. 1.7e-26;
Matches 272; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 964 TTTTGTGTTGTTGACCGAGTCTTACTCTGTCAACCCAGGCTGAGTGTATGTCATG 1023
Db 513 TTTTCAATTTTGTGAGCAGAGTCTCACTCTGCTCACTCAGCTGGAGTGAATGCAATG 454
QY 1024 GTCTCAGTCTACTGCAACCTCCGCTCCGGGTTCAAGCATTTCTCCACCTCAGCTCC 1083
Db 453 GTTTCAGCTCACTGCAACCTCTGCTCTCCGGTTCAAGTGGTTCTCCACCTCAGCTCC 394
QY 1084 CAGTAGCTGGGACTACAGCATGAGCCAGCCGCGCTAATTTTGTATTTTAGTA 1143
Db 393 CGAGTAGCCAGGACTACAGCATGAGCCAGCCGCGCTAATTTTGTATTTTAGTA 334
QY 1144 GAGACGTGGTTCCACTATGTTGCCAGGCTGATCTCGAACTCTGACCTTGTATCCACC 1203
Db 333 GAGATTGATTTCATCTGTTGCCAGGCTGGTCTTGAACCTCTGACCTCGTATCCACC 274
QY 1204 CGCTCGGCTGCCAAAGTCTGGGATTACAGCGTGAACCTTGTGCTTGGCCGATTTT 1263
Db 273 TGCTCGGCTGCCAAAGTCTGGGATTACAGCGTGAACCTTGTGCTTGGCTTACAGT 214
QY 1264 TTAATAAATGTTTCTTATGTCAGTTT 1290
Db 213 TTGTGTTTAAATCTTTCACGTTTTTTT 187

RESULT 3

AQ745353/c
LOCUS
DEFINITION
DKFZp434B117_r1_434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION
DKFZp434B117.5
VERSION
AL041450.1 GI:5420801
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 538)
Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wienann, S.
AUTHORS
EST (Bioecker, et al.)
TITLE
Unpublished (1999)
JOURNAL
Contact: MIPS
COMMENT
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

```
/mole="vector: p1r1p1ex2; site_1: s111A; site_2: s111B"
```

RESULT 5	ACCESSION	REFERENCE
CD558243/c	VERSION	AUTHORS
LOCUS	KEYWORDS	
DEFINITION	SOURCE	
	ORGANISM	

CD558243 1030 bp mRNA linear EST 11-JUN-2003
 AGENCOURT 14365567 NIH_MGC 181 Homo sapiens cDNA clone
 IMAGE:3035963 5', mRNA sequence.
 CD558243
 CD558243.1 GI:31584311
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Plrimates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1030)
 NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgaabs-r@mail.nih.gov Tissue Procurement: Dr. Michael Brownstein cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: NDAM476 row: o column: 12 High quality sequence stop: 564.

```

FEATURES
source
high quality sequence map: 50x.
Location/Qualifiers
1. .1030
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30395963"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances )"
/clone_lib="NIH MGC 181"
/note="vector: pCMV-SPORT6.1; Site 1: NotI; Site_2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen) Note: this is a NIH MGC library."

```

ORIGIN	Query Match	9.23;	Score 234.6;	DB 14;	Length 1030;
	Best Local Similarity	62.94;	Pred. No. 7.2e-26;		
	Matches	471;	Conservative	0; Mismatches 254;	Indels 24; Gaps 6;

645	QY	TAATTTCAAATATCTCACTTGCCCTCTTAAATAGAGTGGGTTTTTAAATATTTTCTGAAGTA	704
780	Db	TAAAAACGAGCTTTTATAGTAGGTTCAATAAAGAGAGATTTTTTACACTTTTTTCGAAAAA	721
705	QY	GGTTTTATGCAATTAATATTTTTTCCCTTTAACTTTCAAACTCAAGGAAAAACCAAGTT	764
720	Db	GCTCCTTTAGCTAGGAATTTTAGGTATTGGAAATTAATACTAAGCCTCCAAAAAATTAGTT	661
765	QY	G-GCCTTGA CTCTGTTTGTGGAATAATTTTAAACTACTGSGTTTAAATTTCTTTTATTTGGTTGT	823
660	Db	GAGAAGTTTATTTCTCTCTTCATTTCTGAAGACTGTTGGTTAGAAATTTTTACAGTGTTTTT	601
824	QY	AATATGACTATTTTTACGTCATATAACAATTTTTTATTTGTTTAAATGACATTTTATTTGTTT	883
600	Db	CCCAAATGTTTGGAAATTCACCTTTTGATCTCTCTGG-----GCCCTGAAGTTTTC	551
884	QY	GTCATATGANAATTTTATGTGCATAGAAACAATTTTTTATTTGCTTGATATATGACTTTATTTGT	943
550	Db	TTTGTAGGAATGTTTTCTATCACAAAGTGAATTTTCTTTGACACATAAAGAATATTCAGT	491
944	QY	TATATGGCTATACAACTAGATTTTTTTGCTGTTTTTGACCCAGATCTTACTCTGTCAACCCA	1003
490	Db	TTTTTTGGGGTTTTTTGTTTTTTTTCCTTTTGTGAGACAGTCTCGCGCTGCACACCCA	431
1004	QY	GGCTGGAGTGTAATGGCAATGTTCTACAGTCTACTGCAACCTCGGCTCTCCGGGGTTCAAGCC	1063
430	Db	GGATGGAGTGCAGGGGTACGATCTGGGCTCACTGCAACCTCCACCTCCACAGATTGAAGCA	371
1064	QY	ATTCTTCCACCTCAGCCTCCCAAGTAGCTGGGACTACAGGCATAGACCCACCGCACCCCGGC	1123
370	Db	ATTCTCTCGCCTCAGCCTCCCGAGTAGCTGGGATACAGGCATGCACCACTACCTGCG	311
1124	QY	TAATTTTTTGATTTTTTAGTAGAGA--CGTGGTTCACATATGTTGGCCAGGCTGATCTCGA	1181
310	Db	TTATTTTTTGATTTTTTAGTAGAGTGGGGGTTTTCCACATGTTGGCCAGGCTGGTCTTGA	251

QY	1182	ACTCTGACCTTGTAATCACACCGCTCGCGCTGGGATTCACAGCGGTGA	1241
Db	250	ACTTCTGACCTTGTAATCACACCTCTCCAGCTTCCAGATGGATTCAGGTGTGA	191
QY	1242	GCCATTGTGCTGGCGGA-----TTTTTTAAAAAATGATTCCTTATGTGTCAGATTTTCATA	1295
Db	190	GCCACGTCGCAGGCCAAGACTATTACGTTTTTAATTTTAAATTTTGTGTCAGATTTTGGAA	131
QY	1296	AGTTTTATTAAAAATGCATTTTCCATTGTGAATGAAGCTTTCAAATTTAT--AGTATAGT	1352
Db	130	AGCTGCATTTTTCCAGGAATTT--GTTTCATCGAAGCATCAAATTTTGTGGTGAAGT	73
QY	1353	TGTTCTAGAAATTTTCTTATCTTTTGGAA	1381
Db	72	TGTTTTAGAAATTCGAATTAATCTTTTTTGGAA	44

RESULT 6	
AI244254	
LOCUS	377 bp mRNA linear EST 02-DEC-1998
DEFINITION	qy90905.x1 NCI CGAP Ut2 Homo sapiens cDNA clone IMAGE:198888 3' similar to contains Alu repetitive element; , mRNA sequence.
ACCESSION	AI244254
VERSION	AI244254.1
KEYWORDS	GI:3839651
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs- email.nih.gov unknown library type Insert Length: 852 Std Error: 0.00 Seq primer: -40UP from Gibco high quality sequence stop: 373

```

FEATURES
source
Location/Qualifiers
1. .377
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1988888"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI-CCAP_Ut2"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: Sall;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.85 kb. Life Technologies catalog #:
11539-012"

```

ORIGIN	Query Match	9.2%: Score 234.4; DB 9; Length 377;
	Best Local Similarity 85.8%; Pred. No. 1.2e-25;	
	Matches 259; Conservative 0; Mismatches 43; Indels 0; Gaps 0;	
Qy	964 TTTTTTTGTTCTTTTGTACCGAGTCTTACTCTGTCCACCGAGCTGGAGTGTAATGGCATG 1023	
Dp	11 TTGTTGTGCTGTTGAGACGAGTCTCACTCTGTCCACCGCTGGAGTGCAACGGCATG 70	
Qy	1024 GTCTCAGCTCATCTGCAACCTCCGGCTCCCGGGTTCAAGCCATTTCTTCCACTCAGCTCC 108	
Dp	71 GTCTCANCTNCTGTAACTCTGCTCCCGGGTTCAAGCGGTTCTCCACCTCAGCTCC 130	
Qy	1084 CAAGTAGCTGGAGCTACAGGCATGAGCCACCGCACCCGGCTAAATTTTGTATTTTGTAGTA 114	
Dp	131 CAAGTAACCTGGACATACAGTGTGTGCCACACACCTGGCTAAATTTTGTATTTTGTAGTA 190	
Qy	1144 GAGACGTGGTTTCCACTATGTTGGCGCAGGCTGANTTCGAATCTCTGACCTGTGTAATCCACC 120	

ORIGIN	Query Match	9.2%	Score 233.2;	DB 9;	Length 599;
	Best Local Similarity 80.5%;		Pred. No. 1.5e-25;		
	Matches 285;	Conservative 0;	Mismatches 68;	Indels 1;	Gaps 1;
QY	913	ATTTTATGCTTGATATGACTTTTATGTATATGGCTATACACTAGATATTTTTGT	972		
Db	355	ACTATCTTCCATTTATTTGATCTCTTTATTTCTTGATTTTTTTTTTTTTTTTT	296		
QY	973	TGTTTTTGACCGAGTCTTATCTC-TGTACCCAGGCTGGAGTGTAAATGGCATGTGCTCAGC	1031		
Db	295	TTTTTTTAGACGGAGTTTCGTCTTTGTCCACCGCCGGAGTGCAATGGCATGATCTCGGC	236		
QY	1032	TCACTGCAACCTCCGCTCCCGGGTTCAAGCCATCTTCCACCTCAGCTCCCAAGTAGC	1091		
Db	235	TCACCGCAACCTCTGCTCCCGGTTCAAGCAATTTCTCTGCTCAGCTTCCCAAGTAGC	176		
QY	1092	TGGGACTACAGGCATGAGCCACCGCACCCGGCTAAATTTTGTATTTTGTAGTAGAGACGTG	1151		
Db	175	TGGGACTACAGGCATGCGCCACACACCTGGCTAAATTTTGTATTTTGTAGTAGAGATGGG	116		

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 1-57, >(TAAA)n#Simple repeat (matched complement)
 62-347, >ALU (matched complement)
 Seq primer: M13 FORWARD
 POLYA=yes

FEATURES

Location/Qualifiers

1..530

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-FLO-bdg-d-16-0-UI"

/tissue_type="Cell lines"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP_FLO"

/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac

(Pharmacia) with a modified polylinker; Site 1: EcoR I;
 Site 2: Not I; NCI CGAP_FLO is a cDNA library derived from
 a pool of mRNA obtained from 4 cell lines from grade III
 chondrosarcoma tissues. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT7T3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GAGGTCGGTG. The cell line
 was provided by Dr James Martin from University of Iowa.
 TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
 TAG LIB=UI-H-FLO
 TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 9.1%; Score 231; DB 14; Length 530;

Best Local Similarity 85.1%; Pred. No. 3.3e-25;

Matches 258; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY	964	TTTTTTTGTGTTTTTACCGAGTCTTACCTGTCAACCGCTGGAGTGAATGGCATG	1023
Db	54	TATTAATATTTTTTAAACACAGTGCTTGTCTGTCAACCGCTGGAGTGAATGGCATG	113
QY	1024	GTCTCAGCTCACTGCAACCTCCGCGCTCCCGGTTCAAGCCATTCTCCACTCAGCCTCC	1083
Db	114	ATCTTGCTCACTGCAACCTCTGCTCCCGGTTCAAGCAACTCTCCACTCAGCCTCC	173
QY	1084	CAAGTAGCTGGACTACAGGATGAGCCACCGCACCGCTGCTTAATTTTGTATTTAGTA	1143
Db	174	CAAGCAGCTGGACTACAGGACACAGCCACCGCTGCTTAATTTTGTATTTAGTA	233
QY	1144	GAGAGTGTTTCCACTATGTGGCCAGCTGATCTCGAACTCTGACCTTGTATTCACC	1203
Db	234	GAGACGGGTTTCACTGTGTGGCCAGGCTGGTCAATGAACCTCTGACCTTGTGACCCACT	293
QY	1204	CGCTCGGCTGCCAAGTGTGAGATTACAGGCTGAGCCATTGTGCTGCGCGATTTT	1263
Db	294	CACCTCGGCTCAACAGTGTGAGATTACAGGCTGAGCCATTGTGCTGCGCGATTTT	353
QY	1264	TTA 1266	
Db	354	TGA 356	

RESULT 15

AW504485

LOCUS

DEFINITION

UT-HF-BNO-alc-h-01-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone

IMAGE:3080665 5', mRNA sequence.

ACCESSION

AW504485

VERSION

AW504485.1 GI:7142152

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 547)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M.B. Soares Lab

cDNA Library Arrayed by: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/brp/image/image.html

The following repetitive elements were found in this cDNA

sequence:

25-105, >ALU 118-139, >(TA)n#Simple_repeat 159-445, >ALU

Seq primer: M13 Forward.

Location/Qualifiers

1..547

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3080665"

/tissue_type="lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

/lab_host="DH10B (LTI)"

/clone_lib="NIH_MGC_50"

/note="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;

Constructed from size fractionated cytoplasmic mRNA

(3.5-4.4kb). Directionally cloned. Cells provided by

Louis M. Staudt, Ph.D. Library preparation by Maria de

Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Query Match 9.1%; Score 230.6; DB 10; Length 547;

Best Local Similarity 75.3%; Pred. No. 3.8e-25;

Matches 287; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY	919	ATTGCTTGATATGATGATTTATTTATATGGCTATACAACTAGATTTTTTTTGTGTTT	978
Db	108	AGTGTGTGAAT	167
QY	979	TGACCGAGTCTTACTCTGTCAACCGAGCTGAGTGAATGGCATGTTCTCAGCTCCTGC	1038
Db	168	AGCGGAGTCTCACCCTGTTGTCCAGCTGGAGTGCAATGGTGTGATCTCGGCTCACTGC	227
QY	1039	AACTCCGCTCCCGGTTCAAGCCATTCTTCACCTCAGCCTCCCAAGTAGTGGGACT	1098
Db	228	AACTCCGCTCCCGGTTCAAGCCATTCTTCACCTCAGCCTCCCAAGTAGTGGGACT	287
QY	1099	ACAGGATAGCACCACCGCTAAATTTTGTATTTTAGTAGAGAGCTGGTTCAC	1158
Db	288	ACAGGCGCCCGCCACACCGCTAAATTTTGTATTTTAGTAGAGAGCTGGTTCAC	347
QY	1159	TATGTTGGCCAGGCTGATCTCGAAGCTTGTACCTGATTAATCCACCGCTCGGCTGCCA	1218
Db	348	CATGTTGGCCAGGCTGATCTCGAAGCTTGTACCTGATTAATCCACCGCTCGGCTGCCA	407
QY	1219	AAGTGTGGGATTACAGGCTGAGCCATTGTGCTGCGCGATTTTTTAAAGATGATTTC	1278
Db	408	AAGTGTGGGATTACAGGCTGAGCCATTGTGCTGCGCGATTTTTTAAAGATGATTTC	1278
QY	1279	TTATGTGAGTTTTCATAAGTT 1299	
Db	468	TGAAGTCATTGGAATGGGGT 488	

us-09-914-651a-2_copy_14040_16580.rst

Mon May 24 08:12:35 2004

Search completed: May 21, 2004, 23:03:24
Job time : 7099.66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 20:49:36 ; Search time 1146.81 Seconds
(without alignments)

10069.828 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcataattttcttttaatt.....cttctctagtgtgaatttt 2541

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	312	12.3	335	16	US-10-431-791-3
2	271.2	10.7	174566	13	US-10-235-192A-37
3	271.2	10.7	174566	15	US-10-020-141-1
4	262	10.3	454	16	US-10-431-791-1
5	252	9.9	157875	10	US-09-935-464-1
6	252	9.9	157875	15	US-10-125-835-1
7	246	9.7	257	16	US-10-431-791-18
8	241.4	9.5	246144	16	US-10-085-117-226
9	239.8	9.4	21470	9	US-09-764-847-1157
10	239.8	9.4	21470	15	US-10-092-154-1157
11	239.6	9.4	133632	13	US-10-087-192-1810
12	239.4	9.4	16086	9	US-09-764-877-2385
13	239.4	9.4	16086	16	US-10-242-515-2385
14	239	9.4	750	13	US-10-027-632-284167

Sequence 284167,
Sequence 7553, Ap
Sequence 7555, Ap
Sequence 3797, Ap
Sequence 3798, Ap
Sequence 3798, Ap
Sequence 316, Ap
Sequence 3, Appli
Sequence 3, Appli
Sequence 256674,
Sequence 256674,
Sequence 226967,
Sequence 226967,
Sequence 127789,
Sequence 127789,
Sequence 214024,
Sequence 214024,
Sequence 3660, Ap
Sequence 3660, Ap
Sequence 574, App
Sequence 27, Appl
Sequence 19, Appl
Sequence 181, App
Sequence 15927, A
Sequence 15927, A
Sequence 253802,
Sequence 253802,
Sequence 1426, Ap
Sequence 27832, A
Sequence 27832, A

750 16 US-10-027-632-284167
9.4 239 16 US-09-764-891-7553
9.4 238.8 10 US-09-764-891-7555
9.4 538 10 US-09-764-891-7555
9.4 15266 9 US-09-764-877-3797
9.4 15266 16 US-10-242-515-3797
9.4 238 9 US-09-764-877-3798
9.4 15271 9 US-10-242-515-3798
9.4 15271 16 US-09-967-768A-316
9.4 143068 9 US-09-804-474A-3
9.4 126512 9 US-10-036-960-3
9.3 62804 14 US-10-027-632-256674
9.3 1877 13 US-10-027-632-256674
9.3 1877 16 US-10-027-632-256674
9.3 656 13 US-10-027-632-226967
9.3 656 16 US-10-027-632-226967
9.3 797 13 US-10-027-632-127789
9.3 797 16 US-10-027-632-127789
9.2 650 13 US-10-027-632-214024
9.2 650 16 US-10-027-632-214024
9.2 13444 9 US-09-764-877-3660
9.2 65973 13 US-10-087-192-574
9.2 176001 17 US-10-210-556-27
9.2 186739 17 US-10-210-556-19
9.2 56423 16 US-10-232-798-181
9.2 539 13 US-10-027-632-15927
9.2 539 16 US-10-027-632-15927
9.2 1741 13 US-10-027-632-253802
9.2 1741 16 US-10-027-632-253802
9.2 40491 13 US-10-087-192-1426
9.2 656 13 US-10-027-632-27832
9.2 656 16 US-10-027-632-27832

ALIGNMENTS

RESULT 1

US-10-431-791-3
; Sequence 3, Application US/10431791
; Publication No. US20030235874A1
; GENERAL INFORMATION:
; APPLICANT: Kao, Chinghai
; APPLICANT: Lee, Sang-Jin
; APPLICANT: Kim, Hong-Sup
; APPLICANT: Lee, KangRyul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; FILE REFERENCE: 1857-ARTI.0222US
; CURRENT APPLICATION NUMBER: US/10431,791
; CURRENT FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-431-791-3

Query Match 12.3%; Score 312; DB 16; Length 335;
Best Local Similarity 99.4%; Pred. No. 1e-44;
Matches 334; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 721 AATTATTTTCTTAACTTCAACTCAAGAAACAGTTGGCTTGACTCTGTTT 780
Db 1 AATTATTTTCTTAACTTCAACTCAAGAAACAGTTGGCTTGACTCTGTTT 60
QY 781 GTGGAATAATTTAACTACTGTTTAACTTTATTTTATTTGTTGTAATATGACTATTATTACG 840

Db	36808	GATAGCTTTTATAGAACTGGAATGCAAGTAAGGCAATAAAAACTCTTCAATCTTAT	36749
Qy	1356	TCCTAGTATTTCTTATCTTTTGTGTAATCTGTCAGGCTGTAGATGCTCTTTTAA	1415
Db	36748	T-----TACTCAITATGTAATAAACAATAATATGCTATATATGACAGTGAACAA	36696
Qy	1416	TAATAATAATATTTTGTGCGCTTTTGTATTTTTTTTCTTATGCTCTGAGAGGGAT	1475
Db	36695	TATAGATATGATGCTGGCTTCTTTTCTTTTGGCAATTTGATAGATCCATGCAIT	36636
Qy	1476	ATGTCAAAATTTACTAGTATCCAAAGATAAACTTTGGCGTTGGCAATCTTTTCTCATC	1535
Db	36635	AGGA-----TTTTCTGCTCTCTTTTATCTAGTTTGGTTTGTGACTTACATTTTTTTC	36582
Qy	1536	TATCTTTGCTTATATTTTATTAATCTGTTCTGTTTGTATTAATGCTCTTTTATCTTC	1595
Db	36581	TTGTTTCACTTTCACTAGTCTCTCTGATCTGTTGTTTATTTCTTTCTGCTGGGT	36522
Qy	1596	TTTGTGTTTACITTTGCTGTTCTTTGTAATAATCTCAGTAGAATGCTTAATTTACAT	1655
Db	36521	TGGGTTTGGTTTCTTCTGTTCTCTAGTACCTTGAGGTGTGACCTTAGATTGCCAATTT	36462
Qy	1656	TCAGTCTTTCTTCAITTTCTACTATGATGATTTAGAGCCATAAATTTCCCTTTTAACTTCC	1715
Db	36461	GTGCTCTTTTACAGACTTTTGTAGTGTCAITTTTATGCTATTAATCTTTCTCTTAA	36407
Qy	1716	CTTTCACCTTCACTACATCTCAAAATTTGGATAGGAGTAGTTTAAATTTATCATTAGTA	1775
Db	36406	--TACTGCTTTTGTCTGTATCCCAAGAGTTTGTAGTTTGTGTCACTATTATTTGTCAGT	36349
Qy	1776	TCTAAATATTTTAAATTTCTGTTATTTCTTTGATCTCTGCAACTATTTTACAAGTATT	1835
Db	36348	TCAARGATTTTAAATTTTCCATCTTGAITTTCAITTTGATGCCAGTGATTTTTCAGGAGC	36289
Qy	1836	TTTTAAATCTCTGAATATATAAGATTTGTTATTTTGTGATCTGATCTCTAAATG	1895
Db	36288	AGATTATTAATTTCCCTGTATTTGTCATAGTTTGGGGTTCTTTTGGAGTTGATTTCC	36229
Qy	1896	AATATA--TTGAGATCAGATAAATGTGGTTTGTAGAGACACTAATCTTTGACAAATTTGTGAG	1954
Db	36228	AGTTTATTTTCACTGTAGCAAGAGTACTTGTCTGTAATTTTAAATTTTCTTAAATTTGTGAG	36169
Qy	1955	GCTTCTTTTGGAACTTAATATGCTCAATTTTATAGAGCTGCTGTGTTCTTTTGGAA	2014
Db	36168	GCTTGTTTTGGCTTACCACATGCTCTATCTTGGAGGATTTCCATGTGCTGATGAATA	36109
Qy	2015	AAACATGTAATTTGATGTTCTTTTGGTTTAAATAATTTTGTATTTGTACATTTAGTTGATTT	2074
Db	36108	GAATTCATAGTCTGCATTTTATTTGGTAAAGATTTCTGTAATTTTCTGTAGATCAATTTT	36049
Qy	2075	GCTTATTTTGGCTGAAATCTCCATTTATCTTAAATGCTCTCTCAITTTGCTGCTTC	2134
Db	36048	TTTTAGAGTATAGTTTAAAGTCCATTTGTTCTTTTGTG-----ACTTTCTGCTTGATG	35996
Qy	2135	CTTTTAAATAGAGATAAATTTTAAATTTATCTCACCTCACTATAGTATGATCTGTTTATA	2194
Db	35995	ACCTGTCTAATGCTGTCAGTTGATTTGAAGCCCCCTACTACTGTGCTGCTCTCA	35936
Qy	2195	TACTATATATAAAATTTTAAATTTCCATAAATTTATTTATGATATAATTTTGGAGACCTA	2254
Db	35935	TTTCTTAGTCTAGTAGTAATCTTTTATAAATTTGGAGGCTCCAGTGTAGGTG--CAT	35978
Qy	2255	TTATCATATAAAACAGAAATTTGATGAAGTATGACAGACTTATATCTTATGATAGCTTT	2314
Db	35877	ATATGTTTAGGATTTACGATATTTTCTGTTGAGACTGCTCCCTTATTTATATATATGCTCCT	35818
Qy	2315	TTTTATCTCGTCATAATGTTATTTGACTTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2374
Db	35817	CTTTGGCTTTTAACTGCTATTTGCTTTAGTTTGTGTTTCTCTCTCTCTCTCTCTCTCTCT	35758
Qy	2375	TTGGTATTTCTTTT 2388	
Db	35757	TCCTGCTTGTCTTTT 35744	

Db	61	GTGGAATAATTTTAAACTACTGTTTAAATTTCTTTTATTTGTTGTAATATGACTATTTTACG	120
Qy	841	TCATATAACAATTTTATTTGTTTAAATGACTTTATTT--GTTTGTCAATGATAATTTT	899
Db	121	TCATATAACAATTTTATTTGTTTAAATGACTTTTATTTGTTTGTCAATGATAATTTT	180
Qy	900	ATGTCATAGAACAATTTTATTTGCTTGATATATGACTTTTATTTGTTATATGCTTATCAAC	959
Db	181	ATGTCATAGAACAATTTTATTTGCTTGATATATGACTTTTATTTGTTAT--TGGCTATACAAC	239
Qy	960	TAGATTTTTTGTGTTTGGCCGAGTCTTACTCTGTCAACCCAGGCTGAGTGAATCG	1019
Db	240	TAGATTTTTTGTGTTTGGCCGAGTCTTACTCTGTCAACCCAGGCTGAGTGAATCG	299
Qy	1020	CATGCTCAGCTCACTGCAACCTCCGCTCCCGGG	1055
Db	300	CATGGTCTCAGCTCACTGCAACCTCCGCTCCCGGG	335
RESULT 2			
US-10-235-192A-37/c			
; Sequence 37, Application US/10235192A			
; Publication No. US2004004389A1			
; GENERAL INFORMATION:			
; APPLICANT: McCarthy, Jeanette			
; TITLE OF INVENTION: Methods and Compositions for Identifying			
; TITLE OF INVENTION: Risk Factors for Abnormal Lipid Levels and the Diseases			
; TITLE OF INVENTION: and Disorders Associated Therewith			
; FILE REFERENCE: MM-011			
; CURRENT APPLICATION NUMBER: US/10/235,192A			
; CURRENT FILING DATE: 2002-09-04			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 37			
; LENGTH: 174566			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-235-192A-37			
Query Match 10.7%; Score 271.2; DB 13; Length 174566;			
Best Local Similarity 53.0%; Pred. No. 1.6e-36;			
Matches 802; Conservative 0; Mismatches 678; Indels 34; Gaps 9;			
Qy	880	GTTTGTCAATGATAATTTTATGTGTCATAGAACAAATTTTATTTGCTTGATATATGACTTTA	939
Db	37228	GTTTGTCAATGATAATTTTATTTATTTGAGTATGTTCTTTTATAGTGAATCAGATTG	37169
Qy	940	TTGTTATATGGCTATA--CAACTAGATTTTGTGTTTGTGACCGAGTCTTACTCTGTC	998
Db	37168	TATATTTATTTTATTTATTTATTTATTTTATTTTATTTTGTGAGATGGAGTCTTGTGTC	37109
Qy	999	ACCCAGCTGAGTGAATGTCATGCTGCTCAGCTCACTGCAACCTCCGCTCCCGGGTTC	1058
Db	37108	ATGAGCTGGAGTGCAGTACAAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTC	37049
Qy	1059	AAGCAATCTTCCACCTCAGCTCCCAAGTAGCTGGGACTACAGCCATGAGCCACCGCAC	1118
Db	37048	AAGGATCTCTCTGCTCAGCTCCCAAGTAACTGGTACTACAGCCATGCAACCACTGC	36989
Qy	1119	COGCT--AATTTTGTATTTTATGAGAGTGGTTCACATATTTGGCCAGGCTGATC	1177
Db	36988	CTGGCTAAATTTTGTATTTTATGATAGATGGGGTTTACCATGTTGGCCAGGCTGATG	36929
Qy	1178	TCGAATCTCTGACCT--TGTAAATCCACCGCTCGGCTGCAAAAGTCTGGGATTACAG	1235
Db	36928	TCGAATCTCTGACCTCAAGTAAATCCGCTCGGCTGCTCCCAAGTGTGGAATTACAG	36869
Qy	1236	CGCTGAGCAATGTGCTGCGGCTTTTAAAAAATGTAATTTCTTATGCTAGTTTTCATA	1295
Db	36868	GTGTGAGCCACTGCACCCAGCCAGATTTTATTTTGTGTAATCAAGGATATGTAATGT	36809
Qy	1296	AGTTTATTTAAATGCAATTTTCAATTTGATGTAAGCTTTCAAAATTTATAGTATAGTTGT	1355


```
; APPLICANT: Lee, KangRyul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 1857-ARTI-0222US
; CURRENT APPLICATION NUMBER: US/10/431,791
; CURRENT FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
; US-10-431-791-1

Query Match 10.3%; Score 262; DB 16; Length 454;
Best Local Similarity 100.0%; Pred. No. 5.6e-36;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 721 AATTATTTTCTTTAACTTTCAAACTCAAGGAAAAACAGTTGGCCCTGACTCTGTTT 780
Db 193 AATTATTTTCTTTAACTTTCAAACTCAAGGAAAAACAGTTGGCCCTGACTCTGTTT 252
QY 781 GTGGAAAAATTTAACTACTGGTTAAATTTCTTTATTTGTTGTAATGACTATTTTACG 840
Db 253 GTGGAAAAATTTAACTACTGGTTAAATTTCTTTATTTGTTGTAATGACTATTTTACG 312
QY 841 TCATATAACAAATTTTATTTGTTTGTAAATGACTTTATTTGTTGTCATATGATAATTTTA 900
Db 313 TCATATAACAAATTTTATTTGTTTGTAAATGACTTTATTTGTTGTCATATGATAATTTTA 372
QY 901 TGTCTAGAACAAATTTTATTTGTTGTCATATGACTTTATTTGTTGTCATATGACTATTA 960
Db 373 TGTCTAGAACAAATTTTATTTGTTGTCATATGACTTTATTTGTTGTCATATGACTATTA 432
QY 961 AGATTTTGTGTTTGTGAC 982
Db 433 AGATTTTGTGTTTGTGAC 454

RESULT 5
US-09-935-464-1
; Sequence 1, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 157875
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-935-464-1

Query Match 9.9%; Score 252; DB 10; Length 157875;
Best Local Similarity 54.0%; Pred. No. 3.4e-33;
Matches 862; Conservative 0; Mismatches 635; Indels 99; Gaps 13;

QY 805 TAATTTCTTTATTTGTTGTAATGACTATTTTACGTCATATAACAATTTTATTTGTTG 864
```

```
Db 33292 TAGGCTTTCTTTGATGGAGACATTTATTATACAGATTTAAATCTCTTTACTCATATTTG 33351
QY 865 TTAATAGACTTTTATTTGTTGTCATATGATAAATTTTATGTCATAGAACAAATTTTATTTGCT 924
Db 33352 TCTGTTTCAGATTTTCTATTTCTTTCTTTGTTCTATCTTGTAGTTGCAAGATACATATGTC 33411
QY 925 TGATATATGACTTTATTTGTTATATATGCTATACAACTAGATTTTATTTGTTGTTTGTACCG 984
Db 33412 CACGAATTTATCAACTCTCTGCTAGGTTTATTTTGTGTTTATTTTGTGATG 33471
QY 985 AGTCTTACTCTGTCACCCAG-----GCTGGAGTGTAAATGGCATGTCTCAGCTCACTGC 1038
Db 33472 AGTCTGACTTTTGTCCAGGCTGGAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 33531
QY 1039 AACCTCCGCTCCCGGGTTCAAGCCATTTCCACCTCAGCCTCCCAAGTAGCTGGACT 1098
Db 33532 AAGCTGGCTCCAGGTTCCAGCCATTTCTCGCTCAGCCTCCGAGTAGCTGGACT 33591
QY 1099 --ACAGGATGAGCCACCCAGCCGGCTAATTTT-----GTATTTTATAGAGAGCTGG 1152
Db 33592 ACACAGGCGCCCGCCACCATGCTGGCTAATTTTGTATTTTATTTTATAGAGATGGGG 33651
QY 1153 TTCCCATATGTTGGCAGGCTGATCTCGAACTCCTGACCTTGTAATCCACCCGCTCGGC 1212
Db 33652 TTTCACTGTTTACAGAGATGTTCTCGATCCCTGACCTCGTGAATCTGCCGCTCGGC 33711
QY 1213 CTGCCAAAGTGTGGGATTTACAGGCGTGAGCCATTTGTGCTGGCCGATTTTAAATAAT 1272
Db 33712 CTCCCAAGTGTGGGATTTACAGGCGTGAGCCACTGCACCCGACGATTTCTGCTAGGTT 33771
QY 1273 GTATCTTATGTCAGTTTTCATAGTTTATTTAAATGCAATTTTCCATTTGATGTAAGC 1332
Db 33772 TTTCAATTAATTTGGCAAT----- 33789
QY 1333 TTTCAATTTATAGTATAGTTGTTCTAGTATTTTCTATCTTTCTTATCTTTCTCAGCG 1392
Db 33790 -----TAGTTGTTCAATAGTATGATCTTATGATCTTTGTTGCAATGT 33833
QY 1393 TCTGTAGATGTCCTCTTTTAAATAAATAATTTATTTGTTGCGGCTTTT--GCTATTTT 1450
Db 33834 TAATTTGATGTCCTCTTTTATTTGATTTTATTTGATTTTATTTGATCTTTTCACTTTT 33893
QY 1451 TTTTCTTATGCTCTTTGAGAGGATATGTCAAATTTTACTAGTGTATCCAAAGATTAAT 1510
Db 33894 TTTCTTGGTTAATCTTTCTAATGTTTGTCAATTTTGTGTTTATCTTTTCAAAAAGCACT 33953
QY 1511 TTGGCGTTGCGCAATCTTTTCTCATCTATCTTTGCTTTATATTTTATTAATCTGTTCTTG 1570
Db 33954 TTTTCTTGTATCTTTTAAATTTTATAGCTCTATATTTGTTGATTTCTGCTCTCACTTTT 34013
QY 1571 TTTTATAATTTGCTCTTTTATCTTTCTTTGTTTACTTTGCTGTTCTTTGTAAAATCCTC 1630
Db 34014 ATTATTTCTCTCTCTCTTTAAATTTTTC--GGATTTTGGTTTCTCGGTTGCTAGGCCCTTG 34072
QY 1631 AGTAGAATGCTTAATTTATGATTTGATCTTCTTTCTTTCTTTCTTCTATCTATGATTTTGA 1690
Db 34073 AAATGCTTTATTTAGGTTGTTTATTTGCAATCTCTCTTCTTATATAATAGAGCCATTTAT 34132
QY 1691 GCCATAAATTTCCCTTTTAACTTTCCCTTTCCACTCAACTACATCTCACAAAATTTGGATT 1750
Db 34133 GCTATTAATTTTCTTTT-----AGACTACTTTTGTGATCTCTATAGTTTGGTA 34186
QY 1751 AGGAGTAGTTTAAATATCATATAGTATCTAAATATTTTAAATCTT--TTCCTTC 1807
Db 34187 TATTGTGATCTATTTTCAATTTTCAAGAAATTTTAAATTTCTCTTCTTAAATTTATTA 34246
QY 1808 TTTGATCCTGCAACTATTTTACAGTATTTTAAATCTT--GAATATAAGATTTGTTATT 1866
Db 34247 ATTGACCCATTTGCTATTTAGGAGCAATGTTGTTAAATTTTCTGTTTATTTATATGTTCCA 34306
QY 1867 GTTATTTGTTGATCTGATCTCTAAATTTGAATATATTTAGATCAGATTAATGTTGTTGTA 1926
Db 34307 AAGTTGTTCTTGGTATTAATTTTATTTTATTTTCCATGATCAGAAAAGATATTTCATA 34366
```


Qy	2045	TATTTTGTATTTGTGACATTAAGTTTGGCTTATTTTGGCTCGAATCTCCATTATC	2104
Db	34487	AGTCTGAAAATGCTATTTTGGTCCAAATTTGGTCTAAAAATGCGAGTTTAAATTCATAGCTTTC	34546
Qy	2105	CTTAATGTGCTCTCTCATTTTGTCTGCTCCCTTTATTAATAGAGATAAATGTTAAATTA	2164
Db	34547	TTTGCTG--ATTGCTGTGTACATGATCTGTCCAATGCTGTGGTGGAGTGTGGAAGTC	34603
Qy	2165	TCTCACCCTCCTATAGTGATGCTGTGTTTTATCTATATATAAAATTTTAAATTTCCCAT	2224
Db	34604	CCCAACTAAATGGAGGCTATCTCTCCATTTAGGTCTCAATCATATTTGCTTTTATATTC	34663
Qy	2225	AAATTTATGTTATGTATAAATTTGGAGAGCTATTATCATATATAACACAGAATGTTGATGAA	2284
Db	34664	TGAGTGTTCAGTCTTAAGGTGTATATATTTTACAACCTGTTA--TATACTCTTGCTAAA	34721
Qy	2285	ATGACAGACTTATACCTTTATGTAGTACGCTTTTTTTTCTCGTCATATGTTATTTGACTTT	2344
Db	34722	TTGGCTCTTTTATTAATATAAATGCTTTCTTTGT-----ATGTTTTAGTTTTT	34772
Qy	2345	GTCCATAAAATTTTTTTTTAAATTAATTTGTTGGTA	2380
Db	34773	GACTTAATGCTATTTTATGTATGTAAAGTATAGCA	34808

```

RESULT 7
US-10-431-791-18
; Sequence 18, Application US/10431791
; Publication NO. US20030235874A1
; GENERAL INFORMATION:
; APPLICANT: Kao, Chinghai
; APPLICANT: Lee, Sang-Jin
; APPLICANT: Kim, Hong-Sup
; APPLICANT: Lee, Kangryul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 1857-ARTI.0222US
; CURRENT APPLICATION NUMBER: US/10/431,791
; CURRENT FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-431-791-18

```

Query Match	9.7%;	Score 246;	DB 16;	Length 257;
Best Local Similarity	99.6%;	Prod. No. 2.6e-33;		
Matches 257;	Conservative	0;	Mismatches 0;	Indels 1; Gaps 1;
QY	721	AAATTATTTTTCCTTTAAACCTTTCAAACCTCAAGGAAAAACCAAGTTGGCCTTGACACTCTGTT	780	
Db	1	AAATTATTTTTCCTTTAAACCTTTCAAACCTCAAGGAAAAACCAAGTTGGCCTTGACACTCTGTT	60	
QY	781	GTGGGAAAAATTTTAAACTACTGTTTTAAATTCCTTTATTCGTTGTAATATGACTATTTTACG	840	
Db	61	GTGGGAAAAATTTTAAACTACTGTTTTAAATTCCTTTATTCGTTGTAATATGACTATTTTACG	120	
QY	841	TCATATAACAATTTTTTATTTGTTTGTAAATGACCTTTATTTGTTTGTGCATATGATAATTTTA	900	
Db	121	TCATATAACAATTTTTTATTTGTTTGTAAATGACCTTTATTTGTTTGTGCATATGATAATTTTA	180	
QY	901	TGTCATAGACAAATTTTTTATTCCTTGATATATGACTTTTATTCGTTTATATGGCTATACAACT	960	
Db	181	TGTCATAGACAAATTTTTTATTCGTTTGATATATGACTTTTATTCGTTTAT-TGGCTATACAACT	239	
QY	961	AGATTTTTTTTGTGTTTT	978	

Db 240 AGATTTTGTGCTTT 257

RESULT 8

US-10-085-117-226

; Sequence 226, Application US/10085117

; Publication No. US20030232334A1

; GENERAL INFORMATION:

; APPLICANT: Morris, David W.

; APPLICANT: Engelhard, Eric K.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER

; FILE REFERENCE: 52945200121

; CURRENT APPLICATION NUMBER: US/10/085,117

; CURRENT FILING DATE: 2002-02-27

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 361

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 226

; LENGTH: 246144

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: variation

; LOCATION: (1)...(246144)

; OTHER INFORMATION: n = any nucleotide

US-10-085-117-226

Query Match 9.5%; Score 241.4; DB 16; Length 246144;

Best Local Similarity 75.3%; Pred. No. 2.8e-31;

Matches 314; Conservative 0; Mismatches 101; Indels 2; Gaps 1;

QY 851 ATTTTATTGTTGTTTAAATGACITTTATTTGTTGTCATATGATAATTTTATGTCATAGAA 910

Db 142504 AATTTTGATGATGCCAATTTATTTTCCTTTGGTTGTTTGGCTTTTGATGTCATAGCT 142563

QY 911 CAATTTTATTGCTTGATATAGCTTTTATGTTTATATGGCTATACAACTAGATTTTTT 970

Db 142564 AAGAAACCATTTGCCCTTAATCCAGGTGATAAAGATTTACATCTATGTTGTTTTTTT 142623

QY 971 GTTGTGTTTG--ACCGAGTCTTACTCTGTCACCCAGGCTGAGTGTATATGGCATGCTCTC 1028

Db 142624 TTTTTTGTGAGACGAGCTCTTGCCCTGTGCGCCAGGCTGGAGTGCATGGTGGGTCTC 142683

QY 1029 AGCTCACTGCACCTCCGCTCCCGGTTCAAGCCATTCTTCCACTAGCCTCCCAAGT 1088

Db 142684 CACTCACTGCACCTCTGCCTCCCAAGTTCAGCAATTTCTTGCCCCAGCCTCCCAAGT 142743

QY 1089 AGCTGGGACTACAGCATGAGCCAGCCAGCCCGGCTAATTTTGTATTTTATTTAGTAGAGAC 1148

Db 142744 AGCTGGGATTACAGCGGTGCATCCACACCTGGGCTAATTTTGTATTTTATTTAGTAGAGAC 142803

QY 1149 GTGGTTTCCACTATGTTGGCAGGCTGATCTCGAACTCTGACCTTGTAACTCACCGGCT 1208

Db 142804 GGGGTTTACCATGTTGGTCAGGCTGGTCTCGAACTCTGATCTGATCCACCGGCT 142863

QY 1209 CGGCTTGCAGAGTGTGCGGATACAGGCGTGAGCCATTTGCTGCTGGCCGATTTTTT 1265

Db 142864 CAGCCTCTCAAAGTGTGGGATTACAGGCGTGAGCCACCGCGCGGCCCATCTTTT 142920

RESULT 9

US-09-764-847-1157

; Sequence 1157, Application US/09764847

; Patent No. US20020132767A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC009

; CURRENT APPLICATION NUMBER: US/09/764,847

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PAM or file wrapper

```

; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1157
; LENGTH: 21470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1157

Query Match
Best Local Similarity 9.4%; Score 239.8; DB 9; Length 21470;
Matches 277; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 964 TTTTGTGTTGTTTGTGACCGAGTCTTACTGTCTGACCCAGGCTGGAGTGAATGGCATG 1023
DB 8153 TTTTGTGTTGTTTGTGACCGAGTCTTACTGTCTGACCCAGGCTGGAGTGAATGGCATG 8212
QY 1024 GTCTCAGTCTCACTCAAGCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCC 1083
DB 8213 GTCTCGGCTCACTCAAGCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCC 8272
QY 1084 CAAGTAGCTGGGACTACAGGATGAGCCACCGCCCGGCTAAATTTTGTATTTTAGTA 1143
DB 8273 TGAGTAGCTGGGACTACAGGATGAGCCACCGCCCGGCTAAATTTTGTATTTTAGTA 8332
QY 1144 GAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTGTAATCCACC 1203
DB 8333 GAGACAGGTTTCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTGTAATCCACC 8392
QY 1204 CGCTCGGCTCCCAAGAGTGTGGGATACAGGCTGAGCCATTGTGCTGGCCGATTTT 1263
DB 8393 TGCTCAGCCTCCCAAGAGTGTGGGATACAGGCTGAGCCATTGTGCTGGCCGATTTT 8452
QY 1264 TTAATAATGATTTCTTATGTCAGTTTTCATAAGTTTGA 1302
DB 8453 GTTTATGTGAAGAGATTTCTCAGTGTATACATCTGTCA 8491

US-10-087-192-1810/c
; Sequence 1810, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1810
; LENGTH: 133632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1810

Query Match
Best Local Similarity 9.4%; Score 239.6; DB 13; Length 133632;
Matches 284; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 909 AACATTTTATGCTTGTGATATAGCTTATTTGTTATATAGGCTATACAACTAGATTTT 968
DB 72842 AGCTATTATTTCTTAAATATATAGCATTCATTTTAAAGGTATATAGGATTTT 72783
QY 969 TTGTTGTTTGTGACCGAGTCTTACTGTCTCACCAGGCTGGAGTGAATGGCATGCTC 1028
DB 72782 TTTTGTGAGACAGAGTCTTGTCTGTGGCCAGGCTGGAGTGAATGGCATGCTC 72723
QY 1029 AGCTCACTGCAACCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCCCAAGT 1088
DB 72722 GGCACACTGCAACCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCCCTGT 72663
QY 1089 AGCTGGAGCTACAGGATGAGCCACCGCCAGGCTAAATTTTGTATTTTAGTAGAGAC 1148
DB 72662 AGCTGGAGTATACAGGACCGCCAGGCTAAATTTTGTATTTTAGTAGAGAT 72603
QY 1149 GTCTGTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTTCTCAGCCTCCCTGT 1208
DB 72602 GGGGTTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTTCTCAGCCTCCCTGT 72543
QY 1209 CGGCTCCCAAGAGTGTGGGATACAGGCTGAGCCATTGTGCTGGCCGATTTTGA 1266
DB 72542 CGGCTCCCAAGAGTGTGGGATACAGGCTGAGCCATTGTGCTGGCCGATTTTGA 72485

RESULT 12
US-09-764-877-2385
; Sequence 2385, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005

; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1157
; LENGTH: 21470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1157

Query Match
Best Local Similarity 9.4%; Score 239.8; DB 15; Length 21470;
Matches 277; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 964 TTTTGTGTTGTTTGTGACCGAGTCTTACTGTCTGACCCAGGCTGGAGTGAATGGCATG 1023
DB 8153 TTTTGTGTTGTTTGTGACCGAGTCTTACTGTCTGACCCAGGCTGGAGTGAATGGCATG 8212
QY 1024 GTCTCAGTCTCACTCAAGCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCC 1083
DB 8213 GTCTCGGCTCACTCAAGCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCC 8272
QY 1084 CAAGTAGCTGGGACTACAGGATGAGCCACCGCCCGGCTAAATTTTGTATTTTAGTA 1143
DB 8273 TGAGTAGCTGGGACTACAGGATGAGCCACCGCCCGGCTAAATTTTGTATTTTAGTA 8332
QY 1144 GAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTGTAATCCACC 1203
DB 8333 GAGACAGGTTTCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTGTAATCCACC 8392
QY 1204 CGCTCGGCTCCCAAGAGTGTGGGATACAGGCTGAGCCATTGTGCTGGCCGATTTT 1263
DB 8393 TGCTCAGCCTCCCAAGAGTGTGGGATACAGGCTGAGCCATTGTGCTGGCCGATTTT 8452
QY 1264 TTAATAATGATTTCTTATGTCAGTTTTCATAAGTTTGA 1302
DB 8453 GTTTATGTGAAGAGATTTCTCAGTGTATACATCTGTCA 8491

US-10-092-154-1157
; Sequence 1157, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1157
; LENGTH: 21470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1157

Query Match
Best Local Similarity 9.4%; Score 239.8; DB 15; Length 21470;
Matches 277; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 964 TTTTGTGTTGTTTGTGACCGAGTCTTACTGTCTGACCCAGGCTGGAGTGAATGGCATG 1023
DB 8153 TTTTGTGTTGTTTGTGACCGAGTCTTACTGTCTGACCCAGGCTGGAGTGAATGGCATG 8212
QY 1024 GTCTCAGTCTCACTCAAGCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCC 1083
DB 8213 GTCTCGGCTCACTCAAGCTCCGCTCCCGGTTCAAGCCATTCTTCCACCTCAGCCTCC 8272
QY 1084 CAAGTAGCTGGGACTACAGGATGAGCCACCGCCCGGCTAAATTTTGTATTTTAGTA 1143
DB 8273 TGAGTAGCTGGGACTACAGGATGAGCCACCGCCCGGCTAAATTTTGTATTTTAGTA 8332
QY 1144 GAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTGTAATCCACC 1203
```

; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2385
; LENGTH: 16086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2385

Query Match 9.4%; Score 239.4; DB 9; Length 16086;
Best Local Similarity 70.8%; Pred. No. 2e-31;
Matches 318; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 817 TGGTTGTAATAGCTATTTTACGTCATATAACAATTTTATGCTATAGAACAAATTTTATGCTTGAATATGACT 876
Db 13032 TGTAAATGTTTGGCTGTGATAGGCAATTTAGTCACTGATGCTGTATTTATTTATTTATGAT 13091

QY 877 ATTGTTTGCATATGATAATTTTATGCTATAGAACAAATTTTATGCTTGAATATGACT 936
Db 13092 ATATTGGTATTATATCTCCCATCACAATTTTGTGCTATTTGTCCCACTCTTGTTTCT 13151

QY 937 TTATTGTTATATGGCTATACAACTAGATTTTTTGTGTTTGTGTTTGTGACCGAGTCTTACTCTG 996
Db 13152 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 13211

QY 997 TCACCCAGGCTGGAGTGTAATGGCATGCTCAGCTCACTGCAACCTCCGCTCCCGGT 1056
Db 13212 TCACCCAGGCTGGAGTGCAATGGCGATCTCGGCTCACTGCAACCTCCGCTCCCGGT 13271

QY 1057 TCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGCTGGGACTACAGGCATGAGCCACCGC 1116
Db 13272 TCATGCCATCTCTCTGCTCAGCTCCCAAGTAGCTGGGACTACAGGCATGAGCCACCAT 13331

QY 1117 ACCCGCTAATTTTGTATTTTATGAGAGACGTGTTTCACTATGTTGGCCAGGCTGAT 1176
Db 13332 GGCGGCTAATTTTGTATTTTATGAGAGACGGGTTTCACTATGTTGGCTAGGCTGT 13391

QY 1177 CTCGAACCTCGACCTTGTAATCCACCGCTCGGCTGCAAGTGCTGGGATTACAGG 1236
Db 13392 CTCGAACCTCGACCTTGATCTGCGGCTTGGCTCCCAAGTGCTGGGATTACAGG 13451

QY 1237 CGTGAGCCATTGCTGCTGGCCGATTTTTT 1265
Db 13452 CGTGAGCCACCGGCTGGCTTTTCT 13480

RESULT 13
US-10-242-515-2385
; Sequence 2385, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496

; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2385
; LENGTH: 16086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-2385

Query Match 9.4%; Score 239.4; DB 16; Length 16086;
Best Local Similarity 70.8%; Pred. No. 2e-31;
Matches 318; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 817 TGGTTGTAATAGCTATTTTACGTCATATAACAATTTTATGCTTGAATATGACT 876
Db 13032 TGTAAATGTTTGGCTGTGATAGGCAATTTAGTCACTGATGCTGTATTTATTTATTTATGAT 13091

QY 877 ATTGTTTGCATATGATAATTTTATGCTATAGAACAAATTTTATGCTTGAATATGACT 936
Db 13092 ATATTGGTATTATATCTCCCATCACAATTTTGTGCTATTTGTCCCACTCTTGTTTCT 13151

QY 937 TTATTGTTATATGGCTATACAACTAGATTTTTTGTGTTTGTGTTTGTGACCGAGTCTTACTCTG 996
Db 13152 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 13211

QY 997 TCACCCAGGCTGGAGTGTAATGGCATGCTCAGCTCACTGCAACCTCCGCTCCCGGT 1056
Db 13212 TCACCCAGGCTGGAGTGCAATGGCGATCTCGGCTCACTGCAACCTCCGCTCCCGGT 13271

QY 1057 TCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGCTGGGACTACAGGCATGAGCCACCGC 1116
Db 13272 TCATGCCATCTCTCTGCTCAGCTCCCAAGTAGCTGGGACTACAGGCATGAGCCACCAT 13331

QY 1117 ACCCGCTAATTTTGTATTTTATGAGAGACGTGTTTCACTATGTTGGCCAGGCTGAT 1176
Db 13332 GGCGGCTAATTTTGTATTTTATGAGAGACGGGTTTCACTATGTTGGCTAGGCTGT 13391

QY 1177 CTCGAACCTCGACCTTGTAATCCACCGCTCGGCTGCAAGTGCTGGGATTACAGG 1236
Db 13392 CTCGAACCTCGACCTTGATCTGCGGCTTGGCTCCCAAGTGCTGGGATTACAGG 13451

QY 1237 CGTGAGCCATTGCTGCTGGCCGATTTTTT 1265
Db 13452 CGTGAGCCACCGGCTGGCTTTTCT 13480

RESULT 14
US-10-027-632-284167/c
; Sequence 284167, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284167
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-284167

Query Match          9.4%; Score 239; DB 13; Length 750;
Best Local Similarity 74.6%; Pred. No. 6.7e-32;
Matches 299; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

QY 865 TTAATGACTTTATTTGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATTTGCT 924
    |||||
Db 451 TAAATTTTATAAGAAATTTCTTTTGCAAAATATCTTTTCATCAATATTTGCTTATATGC 392
    |||||
QY 925 TGATATATGACTTTTATTTGTTTATATGCTATATACAACTAGATTTTGTGTTTGTGACCG 984
    |||||
Db 391 TTGATGCTTATCAATTTGTTGTTGCTCTATCTCAATTTTATTTTATCTTTTGTGAGACGG 332
    |||||
QY 985 AGTCTTACTCTGTCAACCCAGGCTGGAGTGTAAATGGCATGGTCTCAGCTCAGTCAACCTC 1044
    |||||
Db 331 AATCTCGCTCTGTCAACCCAGTTTGCAGTGCAGTGCATGATCTCGGCTCACTGCAACCTC 272
    |||||
QY 1045 CGCCTCCCGGGTTCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGTGGGACTACAGGC 1104
    |||||
Db 271 CGCCTCCCGAGTTCAAGCAATTTCTTGCCTCAGCTCCCAAGTAGTGGGACTACAGGC 212
    |||||
QY 1105 ATGAGCCACCGCACCCGGCTAAATTTTGTATTTTATGATAGAGACGTGTTCCACTATGTT 1164
    |||||
Db 211 AAGTGCCACCACTCTGGCTAAATTTTGTATTTTATGATAGAGACAGGGTTTCCACTACT 152
    |||||
QY 1165 GGCAGGCTGATCTCGAACTCTCTGACCTTGTAAATCCACCCGCTCGGCTGCGCAAGTGC 1224
    |||||
Db 151 GGCAGGCTGCTCTTGAACCTCATGACCTCGTGAATCCACCTGCTGCGCTGCCCAAGTGC 92
    |||||
QY 1225 TGGATTACAGGCGTGAGCCATTGTGCTGCGCGATTTT 1265
    |||||
Db 91 TGGATTACAGGCGTGAGCCACCTGCGCCAGCTACTTTTAT 51

Search completed: May 22, 2004, 02:28:16
Job time : 1154.81 secs
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284167
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-284167

Query Match          9.4%; Score 239; DB 13; Length 750;
Best Local Similarity 74.6%; Pred. No. 6.7e-32;
Matches 299; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

QY 865 TTAATGACTTTATTTGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATTTGCT 924
    |||||
Db 451 TAAATTTTATAAGAAATTTCTTTTGCAAAATATCTTTTCATCAATATTTGCTTATATGC 392
    |||||
QY 925 TGATATATGACTTTTATTTGTTTATATGCTATATACAACTAGATTTTGTGTTTGTGACCG 984
    |||||
Db 391 TTGATGCTTATCAATTTGTTGTTGCTCTATCTCAATTTTATTTTATCTTTTGTGAGACGG 332
    |||||
QY 985 AGTCTTACTCTGTCAACCCAGGCTGGAGTGTAAATGGCATGGTCTCAGCTCAGTCAACCTC 1044
    |||||
Db 331 AATCTCGCTCTGTCAACCCAGTTTGCAGTGCAGTGCATGATCTCGGCTCACTGCAACCTC 272
    |||||
QY 1045 CGCCTCCCGGGTTCAAGCCATTCTTCCACCTCAGCTCCCAAGTAGTGGGACTACAGGC 1104
    |||||
Db 271 CGCCTCCCGAGTTCAAGCAATTTCTTGCCTCAGCTCCCAAGTAGTGGGACTACAGGC 212
    |||||
QY 1105 ATGAGCCACCGCACCCGGCTAAATTTTGTATTTTATGATAGAGACGTGTTCCACTATGTT 1164
    |||||
Db 211 AAGTGCCACCACTCTGGCTAAATTTTGTATTTTATGATAGAGACAGGGTTTCCACTACT 152
    |||||
QY 1165 GGCAGGCTGATCTCGAACTCTCTGACCTTGTAAATCCACCCGCTCGGCTGCGCAAGTGC 1224
    |||||
Db 151 GGCAGGCTGCTCTTGAACCTCATGACCTCGTGAATCCACCTGCTGCGCTGCCCAAGTGC 92
    |||||
QY 1225 TGGATTACAGGCGTGAGCCATTGTGCTGCGCGATTTT 1265
    |||||
Db 91 TGGATTACAGGCGTGAGCCACCTGCGCCAGCTACTTTTAT 51

RESULT 15
US-10-027-632-284167/c
; Sequence 284167, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284167
; LENGTH: 750
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 17:11:06 ; Search time 196.141 Seconds
(without alignments)
7189.364 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcataattttttttaat.....cttctctagtgtgaatttt 2541

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	244.2	9.6	9573	4	US-09-220-132-168
C 2	238.2	9.4	43795	3	US-08-742-185-101
C 3	236.4	9.3	62804	4	US-03-800-960-3
C 4	236.4	9.3	62804	4	US-10-096-960-3
C 5	232.8	9.2	72604	4	US-09-268-992-7
C 6	232.8	9.2	72604	4	US-09-657-474-7
C 7	232.6	9.2	74962	4	US-09-685-853A-3
C 8	228.8	9.0	1701	3	US-09-078-294-9
C 9	228.8	9.0	6799	4	US-09-620-312D-299
C 10	227.6	9.0	28001	4	US-09-819-993-3
C 11	227.6	9.0	28001	4	US-10-193-295-3
C 12	227.4	8.9	80246	3	US-09-078-294-4
C 13	227.2	8.9	50000	4	US-03-146-053-3
C 14	226.5	8.9	63000	4	US-03-780-172-18
C 15	226.2	8.9	46718	4	US-09-816-093-3
C 16	225.8	8.9	80595	3	US-09-078-294-3
C 17	225.6	8.9	152331	3	US-09-128-155-16
C 18	225.2	8.9	43950	4	US-09-735-934A-3
C 19	225.2	8.9	43950	4	US-10-060-332-3
C 20	223.4	8.8	63000	4	US-09-780-172-18
C 21	223.2	8.8	63588	4	US-09-873-404-3
C 22	223	8.8	955	4	US-09-620-312D-228
C 23	222	8.7	14747	4	US-09-608-285A-42
C 24	222	8.7	14747	4	US-09-557-800C-42
C 25	222	8.7	15977	4	US-09-608-285A-59
C 26	221.8	8.7	64467	4	US-09-803-671B-3
C 27	221.4	8.7	90541	4	US-09-759-359A-3

C 28	221.2	8.7	84495	4	US-09-797-906-3	Sequence 3, Appli
C 29	221.2	8.7	116592	4	US-09-818-512-3	Sequence 3, Appli
C 30	221	8.7	43069	4	US-09-292-542A-1	Sequence 3, Appli
C 31	221	8.7	116592	4	US-09-818-512-3	Sequence 3, Appli
C 32	221	8.7	162450	4	US-09-345-882-1	Sequence 1, Appli
C 33	220.8	8.7	169998	4	US-09-676-610B-24	Sequence 24, Appli
C 34	220.8	8.7	197496	4	US-09-877-177A-10	Sequence 10, Appli
C 35	220.4	8.7	4444	4	US-09-879-833-3	Sequence 3, Appli
C 36	219.8	8.7	59065	4	US-09-813-817-3	Sequence 3, Appli
C 37	219.8	8.7	59065	4	US-09-978-197-3	Sequence 3, Appli
C 38	219.6	8.6	2886	2	US-08-687-080-55	Sequence 55, Appli
C 39	219.2	8.6	38564	4	US-09-734-673-3	Sequence 3, Appli
C 40	219.2	8.6	75395	4	US-09-984-890-3	Sequence 3, Appli
C 41	218.8	8.6	246240	2	US-08-724-394A-20	Sequence 20, Appli
C 42	218.8	8.6	246240	2	US-08-724-394A-21	Sequence 21, Appli
C 43	218.8	8.6	246240	2	US-08-724-394A-22	Sequence 22, Appli
C 44	218.4	8.6	28664	4	US-09-564-805-28	Sequence 28, Appli
C 45	218.2	8.6	81001	4	US-09-750-580-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-220-132-168/c

; Sequence 168, Application US/09220132

; Patent No. 6506607

; GENERAL INFORMATION:

; APPLICANT: Shyjan, Andrew W.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT

; FILE REFERENCE: 07334-074001

; CURRENT APPLICATION NUMBER: US/09/220.132

; CURRENT FILING DATE: 1998-12-23

; PRIOR APPLICATION NUMBER: US 60/079,303

; PRIOR FILING DATE: 1998-03-25

; PRIOR APPLICATION NUMBER: US 60/068,821

; NUMBER OF SEQ ID NOS: 191

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 168

; LENGTH: 9573

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-220-132-168

Query Match 9.6%; Score 244.2; DB 4; Length 9573;

Best Local Similarity 53.5%; Pred. No. 2.3e-38;

Matches 807; Conservative 0; Mismatches 648; Indels 54; Gaps 12;

Qy	843	ATATAACAATTTTATGTTTAAATGACTTTATTTGTTGTCATATGATAATTTATG	902
Db	5065	AAATAACGTTTCTTTTATTCACCTCATTTGGTTGATATTTACTGTTTGCATATTTT	5006
Qy	903	TCATAGAACAAATTTTATGCTTGATATGACTTTATTTGTTATGCTATACCACTAG	962
Db	5005	TGCAATTTGAAATGTACTGCTATAACATGTGCAAGTGTCTTTTCTTTCTTTT	4946
Qy	963	ATTTTGTGTTGTTT-----GACCGAGTCTTACTC-TGTCACCCAGGCTGGAGTGA	1016
Db	4945	TTTTTTTTTTTTTTTTTTTGGAGACAGAAATTCGCTCTGTTGCCAGGCTGGAGTCAA	4886
Qy	1017	TGCGATGCTCTAGCTCAGTCACTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG	1076
Db	4885	TGTTGCAATCTTGCTCAGTCACTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG	4826
Qy	1077	AGCTCCCAAGTAGCTGGGACTACAGCATGAGCCACCGCCGCTAATTTTGTATT	1136
Db	4825	AGCTCCAGTAGTAGGGGATTAACGATGACCATGTCCAGCTAA-TTTTGTATT	4767
Qy	1137	TTTAGTAGACGTTGGTTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCT-TGT	1195
Db	4766	TTTAGTAGACGTTGGTTTCAACATGTTGGCCAGGCTGGTCTTGTGAACTCTTACCTCAGA	4707

Qy	1196	AAATCCACCCGCTCGGCTCCAAAGTGTGGGATTACAGGCGGTGAGCCATTGTGCTGG	1255
Db	4706	TGATCCGCGCCGCTATCTCCCAAGTGTGGGATTACAGGCGATGACCGTGC	4647
Qy	1256	CGGATTTTAAAAAATGTATCTTATGTGCAGTGTTCATAGTGTTTTATTTAAATGCATT	1315
Db	4646	CC-----ATACAAGTGTCTTTTCAAATAATAACTTATTTTCCAGGAATTTATCCAT	4595
Qy	1316	TTCCATTTGATGTAAAGCTTTCAAATTTATAGTATAGTGTTCCTAGTATTTTCTTATCTT	1375
Db	4594	CTCTTCAGATTTTCTAGTCTGTGTGCATTAAGGGTGTCTATAGTAG-CCTTGAAATCTT	4536
Qy	1376	TTGTAAATCTGTTCAGCGTCTGTAGATGCGCTCTTTTAAATAAATAATATTTGTGTG	1435
Db	4535	TCATATTTCTGTGTGTAATGGTGTGAATATTTCCCATTTCACTTCTAATTCAGCTTATTTG	4476
Qy	1436	CGCTTTTGTCTATTTTTTCTTATTCCTTCGTGAGAGGGATATGCAAAATTTACTAGTGA	1495
Db	4475	GATCGTCTCCCTTTTTTCTTGTGTAATCTCACTAACGGTCTGTGCAATTTGTATCTC	4416
Qy	1496	TCCAAAGAAATAAATCTTGGCGTTGGCAATCTTTTCTCACTATCTTGTCTTTATATTTTA	1555
Db	4415	TTCAAATAACAGTTTTGTTCATTTATCTTTTGATTTTTTGTTCAAATTTAT	4360
Qy	1556	TTAATCTGTTCCTGTGTTTATAAATGCGCTCTTTTATCTCTTGTGTGTTACTTTCGTGT	1615
Db	4359	TTAGTCTGTCTCGATCTTGTATTTCTTTCTTCTCTGGGTTCAGTTCGTGTGTTCTT	4300
Qy	1616	CTTTGTAAATCCTCAGTAGAATGCTTAACTTATTGACATTCAGTCTTCTTCATTTCTA	1675
Db	4299	GTTTCTTTGAGT-----TCCTTGAGGTGCACCTTAGATGTCTGTTCATGCTGT	4251
Qy	1676	CTATGAGTATTAGAGCCATAAAATTCGCCCTTTAACTTCCTTCCCTTCCACTTCAACTAC	1735
Db	4250	TTCAGACTATTTAATGCTATGAACGTCTCTTAGTTTTGCTGT-----ATC	4204
Qy	1736	TCACAAAATTGGATTAGAGTAGTTTAAATATCAATTAGTATCTAAATATTTTAAATTC	1795
Db	4203	TCAGAGGTTTTAAATAGTGTGTGCACATATTCTTTCAGTTTCAAGCAATTTTAAATTC	4144
Qy	1796	TGTAATTTCTCTTTGATCCGTGA-----ACTATTTACAGATTTTCTTAAATCTCTGA	1849
Db	4143	CTCTCTGATTTCAATGTTTAAACCAAGGTCAATTCAGGACAGATTAATTTTCAATG	4084
Qy	1850	ATATAAGATTTGTTATTTGTTATTTGTTGATCTGATCTCTAAATGGAATATATTGAGATC	1909
Db	4083	ATTGTATAGTTTAAAGGTTCTTTTGGAGTTAATTTCCAGTTTTATTCATTTGTGCTC	4024
Qy	1910	AGATAATGTGGTTGTAGGACATAAATCCTTTGACAAATGTTGAGGCTTCTTTTGGAAAC	1969
Db	4023	TGAAAAGGTACTTGACATAAATTTCCAAATTTTCTTAAATTTGTGTGAGACTGTGTTGTG	3964
Qy	1970	TAATATGTCTCAATTTTATATAGAGTTCCTGTGTTTCTTTGGGAAAAACAATGATTTGAT	2029
Db	3963	TATCTATGATTTATCTTGGAGAACGTTCTATGTGCTGATTTAAATATGATTAATCTCGA	3904
Qy	2030	GGTGTGTTGGTTAATATTTTGTATTTGTACATTTAGTTGTAGTTTGTCTTATTTTGGCT	2089
Db	3903	AGTTGTGGGTAGAATGTTAGGTAAATATCTGCTAAATCTATTTGTTCTAGGATATAATT	3844
Qy	2090	GAAATCTCCATATCCCTTAATGTGCTCTCTCAATTTTGTCTGCTCCCTTTATTAATTAG	2149
Db	3843	AAAGTCAAAATTTTCTTTATTGACTTTCTGCTTTGATGTCTGTCTAGTGTCTGCCATGG	3784
Qy	2150	ATAAATGTTAAATATCTCACTCACTATAGTATGCTGTTTTTATACTATATATATAAA	2209
Db	3793	AGTATTG--AGTCCCTACTATTTATGCTGTGCGCACTATCCCATTTCTTAGTCTAGT	3726
Qy	2210	ATTTAATAATCAAAATTTATGTTATGTTATAATTTGGAGACCTATTATACATATATAAAC	2269
Db	3725	AATAATCTGTTTATAAATTTTGGGAGCTCCAGTATTAGTGTCATATATATTTAGGATGTG	3666

QY	2270	AGAAAT-TGTTGATCAAAATGACAGACACTATATACCTTTATGTAGTAGGCTTTTTTATCTCTGTCAT	2322
Db	3665	ATATTGTCTGTTGGACCGATCCTTTTATCATTTATAAAATTTCCCTCTTTTGTCTTTTATA	3606
QY	2329	AATGTTATT	2337
Db	3605	ACTGTTGT	3597
RESULT 2			
US-08-742-185-101			
; Sequence 101, Application US/08742185			
; Patent No. 6020476			
; GENERAL INFORMATION:			
; APPLICANT: Page, David C.			
; APPLICANT: Reijo, Renee			
; APPLICANT: Saxena, Richa			
; APPLICANT: Hawkins, Trevor			
; APPLICANT: Reeve, Mary Pat			
; TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA			
; NUMBER OF SEQUENCES: 102			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.			
; STREET: Two Militia Drive			
; CITY: Lexington			
; STATE: Massachusetts			
; COUNTRY: US			
; ZIP: 02173			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/742,185			
; FILING DATE: 30-OCT-1996			
; CLASSIFICATION: 435			
; PRIORITY APPLICATION DATA:			
; APPLICATION NUMBER: US 08/690,734			
; FILING DATE: 31-JUL-1996			
; PRIORITY APPLICATION DATA:			
; APPLICATION NUMBER: US 08/310,429			
; FILING DATE: 22-SEP-1994			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Granahan, Patricia			
; REGISTRATION NUMBER: 32,227			
; REFERENCE/DOCKET NUMBER: WHI94-07A2			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (617) 861-6240			
; TELEFAX: (617) 861-9540			
; INFORMATION FOR SEQ ID NO: 101:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 43795 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: DNA (genomic)			
US-08-742-185-101			
Query Match 9.4%; Score 238.2; DB 3; Length 43795;			
Best Local Similarity 50.4%; Pred. No. 4.3e-37;			
Matches 775; Conservative 0; Mismatches 743; Indels 21; Gaps 7;			
QY	912	AAATTTTATTCCTTGATATATGACTTTTATCTTATATGGCTATACAACTAGATTTTTTTCG	971
Db	16692	ACTCATATGTCGTCTGTCAGGTTTCTATTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1675
QY	972	TGTTTTTGACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGTAATGCGATGGTCTCAGC	1031
Db	16752	TTTTTTGAGAGGAACTCTCTCTGTCAACCAGGCTGAGTGAGTGTCATCTCGGC	1681
QY	1032	TCACTGCAACTCGCGCTCCGGGTTTCAAGCCATTCTTCCACCTCAGCCTCCCAAGTAGC	1091

Db 16812 TCACTGCAAACTCTCTCTCGGGTTCATGCGAGTCTCTCTGCTCAGACTCTCTGAGTAGC 16871
Qy 1092 TGGGACTACAGGCACTGAGCCACCGCAGCGGCTAA--TTTTGTATTTTGTAGTAGAGC 1149
Db 16872 TGGGACTACGGGCACTCTCATGATGCCAGCTAATTTTGTATTTTGTAGTAGAGC 16931
Qy 1150 TGGTTCACACTATGTTGGCAGGCTGATCTCGAATCTCTGACCTTGTATTCACACCGGCTC 1209
Db 16932 GGGTTTCACCGTGTAGCAGGATGCTCTCATCTCTGACCTCATATTACCGCCGCGCTC 16991
Qy 1210 GGCCTGCCAAAGTCTGGGATTTACAGGCTGAGCCATGTCGCTGCGCGGATTTTAA 1269
Db 16992 TGCCTCTCAAGTCTGGGATTTACGCTGTGAGCCAGCGTGGCGCGCGGTTT---C 17048
Qy 1270 AATGTATTTCTATGTCAGTTTCTATAGTTTCTATATTTTAAATGCA--TTTTCATTTGATG 1327
Db 17049 TATTTCTTACTGGTTCATCTTGGTAGGTTTATGTCAGGAATTTGTCTATTTCTTT 17108
Qy 1328 TAAGCTTTCAAAATTATAGTATAGTTTCTAGTAATTTCTTATCTTTGTATATCTGTT 1387
Db 17109 CTGCTAGGTTCTCTGAATTTTGGTGTATAGTTGTTCAGAAATAGTCTGTAAATGATCTTT 17168
Qy 1388 CAGCGTCTGTAGATGCTCTTTTAAATAAATAATTTTGTGGCTTTTGTCTAT 1447
Db 17169 ATATTTCTCTGAGTCTCTATTTCAATCTCGATTTTATTTATTTAGTCTTTTCTCT 17228
Qy 1448 TTTTCTCTTATGCTCTGAGAGGATGATGCAAAATTTACTAGTATGTAACCAAGATPAA 1507
Db 17229 TTTATTTCTGTTGTGTAGCTAACCGTTTCAAAATTTTGTACCTTTTCAAAAACCA 17288
Qy 1508 ACTTTGGGTTGGCAATCTTTTCACTATCTTTGCTTTTATATTTATTTATTTCTGTTTC 1567
Db 17289 ACTTTTCTTTTGTGATCTCTGATTTTATATGTCGCCATTTTATTTCTGCTGATC 17348
Qy 1568 TTGTTTATAATGCTCTTTTATCTTCTTGTGTTTACTTTGCTGTTTGTGTAATC 1627
Db 17349 TTCAATATGTTTCTTCTACTAAT--TTGGTCAAGTTTGTCTTACTTTTCTAGTTC 17406
Qy 1628 CTCAGTAGAATGCTTAACTATTGACATTCAGTCTTTCTTCAATTTCTACTATGATATT 1687
Db 17407 CTTGATATGATCGTTAGTGTTTTAAATTT--TATTTCCACTTTTCGATACAGTATT 17463
Qy 1688 AGAGCCATAATTTCCCTTTAACTTCCCTTTTCACTTCACTCACTCACTCAAAATGG 1747
Db 17464 ATTGCTGTAAGTTTCTCTTAAATGCGCTTTGTTGATCGATAGGTTTCAAAATGTTG 17523
Qy 1748 ATTAGGAGTAGTTAAATATCATTAGTATCTAAATATTTTAAATTTTCTGATTTCTTC 1807
Db 17524 CTGTTTCAATTTGTTCCAGAAATTTTAAATATGCTTTTAAATTTTCTATTTGACTTC 17583
Qy 1808 TTTGATCCTGCAACTATTTCAGATTTTAAATCTCTGAATATAAAGATTGTTATG 1867
Db 17584 TCAITCAGGAGCAGTTGTTTAAATTTCCATGATTTTGTGTAATTTCAAAAGTTACTCTG 17643
Qy 1868 TTATTTGTTGATCTGATCTCTAAATGAAATATGATGATGATGATGATGATGATGATG 1927
Db 17644 TTATGCAATTTCTAGTTTATTTCCATTTGATTAAGAGAAGATATCTTACCATGATT--CAG 17702
Qy 1928 GACACTAATCTTTGACATTTGTTGAGCTTCTTTTGAACCTTAATATGCTCAATTTT 1987
Db 17703 TTCAATTAATTTGTTGAACCTTAATTTGTTGGAAGAATGTTGGTCTGTCCCCAAGAAC 17762
Qy 1988 TATAGAGTTCTGTGTTTCTTTGGGAAAAAATGATTTTATGATGTTGTTGTTGTTTAAAT 2047
Db 17763 AGTCCATGTCAGAAATTTACTGATGAAGAAGATGTTATTTCTGAGCTATTTGATGAAT 17822
Qy 2048 TTTGATTTGATATGTTGATTTGCTTATTTATTTTGGCTGAAATCTCCATTTATCTT 2107
Db 17823 GTTCTAGAAATGCTGTTAGTCTGTCTGGTCTCAGATTCATTTCTGTTTAAATGTTTCT 17882
Qy 2108 AATGCTCTCTCATTTTGTCTGCTCTCTTTTAAATTTAGATGAATGTTAAATTTATCT 2167
Db 17883 TTGTTGTTTCTGTATCTGTCTAA 17942

Qy 2168 CACCTCACTATAGTCACTGTTTATTAATATATATATATATATATATATATATATATAT 2227
Db 17943 TGCAGAGAGTGAAGTGTGGAAGTCCCAACTATTAATTAATTAATTAATTAATTAAT 18002
Qy 2228 TTATGTTATGATTAATTTGGAGACCTA-----TTATCATATAATAAACAAGATTTG 2279
Db 18003 TAGATTTAATAATAATTTGGTTTTTATATAATCTGGGTGCTCTGATGTTGGGTGAATATA 18062
Qy 2280 ATGAATGACAGACTTATACATTAATGATAGTCCCTTTTATCTCGTCAATGTTATTTG 2339
Db 18063 TTCACAAATGTTATATCTCTGGCTCAATGGACCCCTTATCATTAATTAATTAATTAAT 18122
Qy 2340 ACTTGTCTCTAAATTTTAAATAATTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 2399
Db 18123 TGGTCTCTTTTACAGTTTTTGAAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 18182
Qy 2400 ATGTCACGTGTTGCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2438
Db 18183 AATCACTTTTGGTTTGGTTTCAATGAATATCTTTT 18221

RESULT 3
US-09-800-960-3
; Sequence 3, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

Query Match 9.3%; Score 236.4; DB 4; Length 62804;
Best Local Similarity 81.7%; Pred. No. 1e-36;
Matches 273; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 928 TATATGACTTTATTTGTTATATGCTATACAACATAGATTTTGTGTTTGTGTTTGTGTTTGT 987
Db 43514 TATATGATATCTTAAAGATGCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 43573
Qy 988 CTTACTCTGACCCAGGCTGGAGTGAATGGATGCTCTCAGCTCACTGCAACCTCCGC 1047
Db 43574 GTCTCACTCTCACCAGGCTGGAGTGAATGGATGCTCTGCTCACTGCAACCTCCGC 43633
Qy 1048 CTCGCGGTTCAAGCAATCTTCCACTCAGCTTCCAGCTCCAGTAGTGGAGTACAGGATG 1107
Db 43634 CTCGCGGTTCAAAATTTCTCAACCTCAGCTTCCAGTAGTGGAGTACAGGCGCG 43693
Qy 1108 AGCCACGCGCCGCGTAAATTTTGTATTTTGTAGTAGAGAGCTGTTTCCATGTTGGC 1167
Db 43694 TGCACACACCGAGCTAAATTTTATATTTTATAGTAGAGAGCGGTTTGTCTATGTTGGC 43753
Qy 1168 CAGGCTGATCTGAACTCTCTGACCTTGAATCAACCGCTCCGCTCCGCTCCGCTCCGCTCC 1227
Db 43754 CAGGCTGCTCTGATCTCTGATCTTGTGATCCACCGCTCCGCTCCGCTCCGCTCCGCTCC 43813
Qy 1228 GATTACAGGCGTGAAGCAATTTGCTGCGCGATT 1261
Db 43814 GATTACAGGCAATGAACCGTCCCGCGCAAT 43847

RESULT 4

US-10-096-960-3
; Sequence 3, Application US/10096960
; Patent No. 6664085
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO0115DIV
; CURRENT APPLICATION NUMBER: US/10/096,960
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 09/800,960
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A, T, C or G
US-10-096-960-3

Query Match 9.3%; Score 236.4; DB 4; Length 62804;
Best Local Similarity 81.7%; Pred. No. 1e-36;
Matches 273; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 928 TATATGACTTTATTTATATGCTATACAACTAGATTTTTTTTGTGTTTTCACCGAGT 987
Db 43514 TATAATGATATCTTAAAGATTCCTTTTCTTTTTTTTTTTTTTTTTTTTTTTAGACGA 43573
QY 988 CTCTACTCTGTCACCCAGGCTGGAGTGTATGGCATGTCTCAGCTACTCAGACCTCCG 1047
Db 43574 GTCTACTCTCACCAGGCTGGAGTGTATGGCATGTCTTGTCTACTCAGACCTCCG 43633
QY 1048 CTCGGGGTTCAGGCATCTTCCACTCAGCTCCCAAGTAGCTGGGACTACAGGCATG 1107
Db 43634 CTCGGGGTTCAGGCATCTTCCACTCAGCTCCCAAGTAGCTGGGACTACAGGCATG 43693
QY 1108 AGCCACCGCACCCGGCTAAATTTTGTATTTTATAGAGAGCTGGTTCCTACTATGTGGC 1167
Db 43694 TGCCACACACCCAGCTAAATTTTATATTTTATAGAGAGCTGGTTCCTACTATGTGGC 43753
QY 1168 CAGGCTGATCTGAACTCTGACCTTGTATCCACCGCTCGCCCTGCCCAAGTGTGG 1227
Db 43754 CAGGCTGGTCTGATCTCTGACCTTGTATCCACCGCTCGCCCTGCCCAAGTGTGG 43813
QY 1228 GATTACAGGCTGAGCCATGTGCTGGCCGATT 1261
Db 43814 GATTACAGGCTGAGCCATGTGCTGGCCGATT 43847

RESULT 5

US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28

; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7
Query Match 9.2%; Score 232.8; DB 4; Length 72604;
Best Local Similarity 68.5%; Pred. No. 5.1e-36;
Matches 318; Conservative 2; Mismatches 144; Indels 0; Gaps 0;
QY 805 TAATTTCTTTATTTGGTTGTAATGACTATTTTACGTATATTAACAAATTTTATTGTTG 864
Db 64262 TTATTTAAATTTTTCGATATTTTAAACGTTTCAATAATGTTATGAGTGTATTAGTCA 64203
QY 865 TTAATGACTTTTATTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATTGCT 924
Db 64202 TAATTTACTTAATCATCTCTCTTTTACACATTAAGATGTTTAAATATATATAT 64143
QY 925 TGATATATGACTTTTATTTATATATGCTATACAACTAGATTTTTTTTGTGTTTTCACCG 984
Db 64142 ATATATATATATATATATATATATATATATTTTTTTTTTTTTTTTTTTTTTAAACAG 64083
QY 985 AGTCTTACTCTGTACACCGGCTGGAGTGTATGGCATGTCTCAGCTACTGCAACCTC 1044
Db 64082 AGTCTTCTCTGTGCGCCAGGCTGGAGTGTGAGTGGCGCATCTCTGCTCACTGCAAGTC 64023
QY 1045 CGCTCCCGGGTTCAGGCCATCTTCCACTCAGCTCCCAAGTAGCTGGGACTACAGGC 1104
Db 64022 TGCTCCCGGGTTCAGGCCATCTCTCTGCTCAGCTCCCGAGTAGCTGGGACTACAGGC 63963
QY 1105 ATGAGCACCGCACCCGGCTAAATTTTGTATTTTATAGTAGAGAGCTGGTTCCTATGTT 1164
Db 63962 GCCCGCACCATGCCCGCTAAATTTTATTTTATAGTAGAGAGCTGGTTCCTGTTT 63903
QY 1165 GGCAGGCTGATCTCGAATCTCTGACTTGTATATCCACCGCTCGCCCTGCCAAAGTGC 1224
Db 63902 ATCCAGGATGCTCTCAATCTCTGACTCTGATCCCGCTCAGCTCCCAAGTGC 63843
QY 1225 TGGGATTACAGGCTGAGCCATGTGCTGGCCGATTTTTTTAAA 1268
Db 63842 TGGGATTACAGGCTGAGCCATGTGCTGGCCGATTTTTTTAAATA 63799
RESULT 6
US-09-657-474-7/c
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/078,044
PRIOR FILING DATE: 1998-03-16
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 72604
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified base
LOCATION: all n positions
OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match 9.2%; Score 232.8; DB 4; Length 72604;
Best Local Similarity 68.5%; Pred. No. 5.1e-36;
Matches 318; Conservative 2; Mismatches 144; Indels 0; Gaps 0;
QY 805 TAAATTCCTTTATGCTGTAATATGACACTATTTTACGTCATATAACAATTTTATGTTG 864
Db 64262 TTAATTAATTTTGAATATTTTAAAGCTTCAATAAATGATTTAGTGTATTAGTCA 64203
QY 865 TTAATGACTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCT 924
Db 64202 TAAATGACTTTATGCTCTGTTTACATTTAGAGATGTTTACATATATATAT 64143
QY 925 TGATATGACTTTATGTTTATATGCTGATACAACTAGATTTTGTGTTTGTGACCG 984
Db 64142 ATATATATATATATATATATATATATATATTTTATTTTATTTTATTTTAAACAG 64083
QY 985 AGTCCTACTCTGTCACCCAGGCTGGAGTGAATGGCATGTTCTAGCTCACTGCAACCTC 1044
Db 64082 AGTCCTGCTCTGTCGCCAGGCTGGAGTGCAGTGGCGCATCTCTGCTCACTGCAAGTTC 64023
QY 1045 CGCCTCCCGGTTCAAGCCATTTCTCCACCTCAGCTCCCAAGTAGCTGGGACTACAGGC 1104
Db 64022 TGCTCCCGGTTCAAGCCATTTCTGCTCAGCTCCCGAGTAGCTGGGACTACAGGC 63963
QY 1105 ATGAGCACCGCACCGGCTAAATTTTGTATTTTATGATAGAGAGCTGTTCCACTATGTT 1164
Db 63962 GCCCGCACCATGCGCGGCTAAATTTTATTTTATTTATGATAGAGAGCTGTTCTCGTGT 63903
QY 1165 GGCAGGCTGATCTCGAAGCTCTGACCTTGAATCCAGCCGCTCGGCTGCGCAAGTGC 1224
Db 63902 ATCCAGGATGCTCTCAATCTCTGACCTCGATCCGCGCCCTCAGCTCCCAAGTGC 63843
QY 1225 TGGGATTACAGGCTGAGCCATTTGCTGCGCGCATTTTATAA 1268
Db 63842 TGGGATTACAGGCTGAGCCACTGCGCGCGCCCAACTTTTATA 63799

RESULT 7
US-09-685-853A-3
Sequence 3, Application US/09685853A
Patent No. 6479270
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
FILE REFERENCE: AND USES THEREOF
FILE REFERENCE: CL000871
CURRENT APPLICATION NUMBER: US/09/685,853A
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: 60/182,194
PRIOR FILING DATE: 2000-02-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 74962
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature

LOCATION: (1)...(74962)
OTHER INFORMATION: n = A,T,C or G
US-09-685-853A-3

Query Match 9.2%; Score 232.6; DB 4; Length 74962;
Best Local Similarity 66.9%; Pred. No. 5.6e-36;
Matches 331; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
QY 769 TTGACTCTGTTTGGAAAAATTTTAAACTACTGTTTAAATTTTCTTTTATGTTGTTGTAATAT 828
Db 52522 TTAATGATTTTGGCCACTTCACTGCTGGCGTGAATCAGTGAATGAAATCTTTTAT 52581
QY 829 GACTATTTTACGTATATAACAATTTTATTTTGTGTTTAAATGACTTTTATTTGTTGTCAT 888
Db 52582 CTATATGTTGTTTATTTTCTTTATTTCCCTTTTGTGTTCTGCAAGTTTTCGT 52641
QY 889 ATGATAATTTTATGTCATAGAACAAATTTTATGCTTGCATATATGACTTTTATGTTATAT 948
Db 52642 ACTTAACATTTTGGTACATAAAAAATTTCAAGTAGTGTGTTTATTTAGTTGACCTGTTT 52701
QY 949 GGCTATACAACACTAGATTTTGTGTTTGTGTTTGTGACCGAGTCTTACTCTGACCCAGGCTG 1008
Db 52702 AAATTCAGGTTTGTGTTGTTGTTGAGACAGAGTCTTGTCTGTGCGCCAGGCTG 52761
QY 1009 GAGTGTAAATGTCATGTTCTCAGCTCACTGCAACCTCCCGGTTTCAAGCCATTTCT 1068
Db 52762 GAGTGCAGTGTGCGATCTCGGCTCACTGCAACCTCTGCTCTGTTCAAGTCAATTTCT 52821
QY 1069 TCCACCTCAGCTCCCAAGTAGCTGGGACTACAGGATGAGCCACCGCAGCTAAAT 1128
Db 52822 CCTGCTCAGCTCCCAAGTAGCTGGGATTTACAGGATGATCACCACCGCCGCTAAT 52881
QY 1129 TTTGTTATTTTAGTAGAGAGCTGTTCCACTATCTGCGCAGGCTGATCTGAACTCTG 1188
Db 52882 TTTGTTATTTTAGTAGAGAGCGGGTTTTCACATGTTGCGCAGGCTGGTCTCGAACTCTG 52941
QY 1189 ACCTTGTAAATCCACCGGCTCGGCTGCAAAAGTGTGGGATTTACAGCGGTGAGCCATTG 1248
Db 52942 ACCTCATGATCTCCACCTCGGCTCCCAAGTGTGGGATTTACAGGTGAGCCACTG 53001
QY 1249 TGCTGCGCGATTTT 1263
Db 53002 TGCTGACAAATTT 53016

RESULT 8
US-09-078-294-9
Sequence 9, Application US/09078294
Patent No. 6265211
GENERAL INFORMATION:
APPLICANT: Choo, Kong-Hong Andy
APPLICANT: Du Sart, Desiree
APPLICANT: Cancilla, Michael R.
TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
FILE REFERENCE: Davies Col
CURRENT APPLICATION NUMBER: US/09/078,294
CURRENT FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 1701
TYPE: DNA
ORGANISM: BAC-F2 contig 5
US-09-078-294-9
Query Match 9.0%; Score 228.8; DB 3; Length 1701;
Best Local Similarity 78.1%; Pred. No. 1.5e-35;
Matches 275; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 914 TTTTATGCTTGAATATGACTTTTATGTTATATGCTATACAACTAGATTTTGT 973
Db 1057 TCTCATTTTGAATTCATTACTCTTCTTTATTTATTTATTTATTTATTTT 1116

QY 974 GTTTTGAACGAGTCTTACTCTGTGTCACCCAGGCTGAGTGTAAATGGCATGGTCTAGTCT 1033
Db 1117 TTCTGAGAAGGAGTCTCGCTCTGTGTGCCAGGCTGGAGTGCAGTGGCGTGAATCTCGGCTC 1176
QY 1034 ACTGCAACTCTCGCTCCCGGGTTCAGCCATTCTTCCACCTCAGCTCCCAAGTAGCTG 1093
Db 1177 ACTGCAACTCTCGCTCCCGGGTTCAGCCATTCTTCCACCTCAGCTCCCAAGTAGCTG 1236
QY 1094 GGACTACAGGCATGAGCCACCGCGGCTAATTTTGTATTTTGTAGTAGACGTGGT 1153
Db 1237 GGAATACAGCATGCGCACCAGCGCGGCTAATTTTGTATTTTGTAGTAGAGTGGGT 1296
QY 1154 TCACATATGTGGCCAGGCTGATCTCGAACTCTGACCTGTGTAATCCACCGGCTCGGCC 1213
Db 1297 TTCACCATGTGTGCTAGGCTGGTGTCAAACTCTGACCTTGTATCCGCGCTGCTCAGCC 1356
QY 1214 TGGCAAGTCTGGGATACAGCGGTGAGCCATTGTGCTGGCGGATTTTT 1265
Db 1357 TCCAAAGTCTGGGATTACAGGTGTGAGCCACCATTGCCCGCGCTACTCTTT 1408

RESULT 9
US-09-620-312D-299/c
; Sequence 299, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_Fl_genes Version 1.0
; SEQ ID NO 299
; LENGTH: 6799
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (249)..(2405)
US-09-620-312D-299

Query Match 9.0%; Score 228.8; DB 4; Length 6799;
Best Local Similarity 81.5%; Pred. No. 2e-35;
Matches 277; Conservative 0; Mismatches 62; Indels 1; Gaps 1;
QY 944 TATATGGCTATACAACTAGATTTTTTTGTGTTTTGACCGAGTCTTACTGTGACACCA 1003
Db 3407 TATATGCATACACTTCTATTTTTTTTTTTTTTTTGTAGACAGAGTCTCAGTCTGTGACACCA 3348
QY 1004 GGCTGGAGTGAATGGAGTGTCTCAGCTCACTCAACCTCCGCTCCCGGTTCAAGCC 1063

Db 3347 GSGTGGAGTGAATGGAGTGTCTCGGCTCACTGCAACCTCCGCTCCAGGTTCAAGCG 3288
QY 1064 ATTCTTCCACTCAGCTCCCAAGTAGCTGGGAGCTACAGCATGAGCCACCGACCCGCGC 1123
Db 3287 ATTCTCTCGCTCAGCTCCTGAGTATCTTGGATTACAGGCATGACACCAACGCGCGCGC 3228
QY 1124 TAA-TTTTGTATTTTGTAGTAGAGGTGTTCCACTATGTTGGCCAGGCTGATCTCGAA 1182
Db 3227 TAAATTTTGTATCTTTTAGTAGAGCGGGTTTCCCATGTTGGCCAGGCTGCTCTTGAA 3168
QY 1183 CTCCTGACCTGTGAATCCACCGGCTCCGCGCTGCCAAAGTGTCTGGGATTACAGCGCTGAG 1242
Db 3167 CTCTGACCTTGTGATCCGCGCTCAGCTCCCAAGTGTCTGAGATTACAGCGCTGAG 3108
QY 1243 CCATGTGCTGCCGCGATTTTTTAAAAATGTATTTCTTAT 1282
Db 3107 CCCCACATGCCGCGCTCTTTTTTTTTTTTTTCTCTCTTT 3068

RESULT 10
US-09-819-993-3
; Sequence 3, Application US/09819993
; Patent No. 6436692
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001195
; CURRENT APPLICATION NUMBER: US/09/819,993
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(28001)
; OTHER INFORMATION: n = A,T,C or G
US-09-819-993-3

Query Match 9.0%; Score 227.6; DB 4; Length 28001;
Best Local Similarity 78.0%; Pred. No. 4.3e-35;
Matches 287; Conservative 0; Mismatches 79; Indels 2; Gaps 1;
QY 909 AACAAATTTTATTTGCTTGATATATGACTTTATTTGTTATATGGCTATACAACTAGATTTT 968
Db 16199 ATCTATGATGAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16258
QY 969 TTGTTGTTTTTGACCGAGTCTTACTCTGTCCACCGAGCTGGAGTGAATGGCATGGTCTC 1028
Db 16259 TTTTTTTTGAGACGGAGTCTGTCTGTCCACCGAGCTGGAGTGAATGATGATGATGATGATGAT 16318
QY 1029 AGCTCACTGCAACTCGGCTCCCGGTTTCAAGCATTTTCCACCTCAGCTCCCAAGT 1088
Db 16319 TGCTCACTGCAAGTCTCTCTCCCGGTTTCAAGCATTTTCCACCTCAGCTCCCAAGT 16378
QY 1089 AGCTGGGACTACAGGCATGAGCCACCGACCCCGCTAA--TTTTTGTATTTTGTAGTAGAG 1146
Db 16379 AGCTGGGCTACAGGCATGAGCCACCGACCCCGCTAA--TTTTTGTATTTTGTAGTAGAG 16438
QY 1147 ACGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCTGACCTTGTATTCACCCCGC 1206
Db 16439 ACGGGGTTTCCACCATGTTAGCCAGGATGTTCTCGATCTCTGACCTTGTATTCGCGCTGC 16498
QY 1207 CTGGGCTGCAAGTCTGGGATTACAGCGGTGAGCCATTGTCCTGGCGGATTTTTTA 1266
Db 16499 CTGAGCTCCCAAGTCTGAGATTACAGGCATGAGCCACCGACCCCGCTTAATTTATT 16558
QY 1267 AAAAATGT 1274

```
Db 16559 AAACTTTT 16566

RESULT 11
US-10-193-295-3
; Sequence 3, Application US/10193295
; Patent No. 6620608
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001195DIV
; CURRENT APPLICATION NUMBER: US/10/193,295
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 08/819,993
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(28001)
; OTHER INFORMATION: n = A,T,C or G
US-10-193-295-3

Query Match
Best Local Similarity 9.0%; Score 227.6; DB 4; Length 28001;
Matches 287; Conservative 0; Mismatches 79; Indels 2; Gaps 1;

QY 909 AACAAATTTTATGCTTGATATATGACTTTATTTGTTATATGCTATACAACTAGATTTT 968
Db 16199 ATCTATGATGATACATGATTGATTGATTGATTGATTGATTGATTGATTGATTGATT 16258
QY 969 TTGTTGTTTGGACCGAGTCTTACTGTGTACCCAGGCTGGAGTGTAAATGGCATGGTCTC 1028
Db 16259 TTTTGTGAGACGGAGTCTGTCTGTACCCAGGCTGGAGTGTAAATGGCATGGTCTC 16318
QY 1029 AGCTCACTGCAACTCCGCTCCCGGTTCAAGCATCTTCCACCTCAGCTCCCAAGT 1088
Db 16319 TGCTCACTGCAAGTCTCTCTCCCGGTTCAAGCATCTTCCACCTCAGCTCCCAAGT 16378
QY 1089 AGCTGGGACTACAGGCTAGCCACCCGCTAA--TTTGTGTTATTTTATGATAG 1146
Db 16379 AGCTGGGCTACAGGCTAGCCACCCGCTAA--TTTGTGTTATTTTATGATAG 16438
QY 1147 ACGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAATCTCTGACCTTGTATCCACCCGC 1206
Db 16439 ACGGGGTTTCAACATGTTAGCCAGGATGTTCTCGATCTCTGACCTTGTATCCGCTGC 16498
QY 1207 CTGGGCTCCCAAGTCTGGGATACAGGCTGAGCATTTGCTGGCGGATTTTAA 1266
Db 16499 CTGAGCTCCCAAGTCTGGGATACAGGCTGAGCATTTGCTGGCGGATTTTAA 16558
QY 1267 AAAAATGT 1274
Db 16559 AAACTTTT 16566

RESULT 12
US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294

; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4

Query Match
Best Local Similarity 8.9%; Score 227.4; DB 3; Length 80246;
Matches 285; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 894 AATTTATGTCATAGAACAAATTTTATTTGTTGATATATGACTTTTATTTGTTATATGGCTA 953
Db 38437 AATGGCATGTAATCAAACTGCTACTTATCTGTTCAATTCATATCTTGAATGTTTATTT 38496
QY 954 TACAACATAGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 1013
Db 38497 TATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 38556
QY 1014 TAATGGCATGCTCAGCTCACTGCAACTCCGCTCCCGGTTCAAGCATTTCTTCAC 1073
Db 38557 CAGTGGGCTGATCTCAGCTCACTGCAACTCCGCTCCCGGTTCAAGCATTTCTTCAC 38616
QY 1074 CTCAGCTCCCAAGTAGCTGGGACTTACAGGCTAGAGCCACCGCACCGGCTAAATTTTGT 1133
Db 38617 CTCAGCTCCCTGAGTAGCTGGGATTCAGAGCTGACCTGACCTGACCTGACCTGACCTG 38676
QY 1134 ATTTTATGATAGAGCGTGGTTTCCATATGTTGGCCAGGCTGATCTCGAATCTCTGACCTT 1193
Db 38677 ATTTTATGATAGAGCGTGGTTTCCATATGTTGGCCAGGCTGATCTCGAATCTCTGACCTT 38736
QY 1194 GTAATCCACCGCTCCGCTCCCGGTTCAAGCATGCTGGGATTCAGGCTGAGCCATTTGCTC 1253
Db 38737 ATGATCCACCGCTCCGCTCCCGGTTCAAGCATGCTGGGATTCAGGCTGAGCCATTTGCTC 38796
QY 1254 GGCCGATTTTATAAAATGT 1274
Db 38797 GGTGCAATGTTTATTTATTT 38817

RESULT 13
US-09-146-053-3
; Sequence 3, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MCG103
; CURRENT APPLICATION NUMBER: US/09/146,053A
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-3

Query Match
Best Local Similarity 8.9%; Score 227.2; DB 4; Length 50000;
Matches 253; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 964 TTTTGTGTTGTTTGTGACCGAGTCTTCTCTCACCAGGCTGGAGTGTATGGCATG 1023
Db 8522 TTTTGTGTTGTTTGTGACCGAGTCTTCTCTCACCAGGCTGGAGTGTATGGCATG 8581
QY 1024 GTCTCAGCTCACTGCAACCTCCGCTCCCGGTTCAAGCATTTCTTCCACCTCAGCTCC 1083
```

Db	8582	ATCTCAGCTCAGTCGAACTCCACCTCCCGGGTTCAAGCGATTTCCCTCGCTCAGCCTCC	8641
Qy	1084	CAAGTAGCTGGGACTACAGCATCAGCCACCGCACCCGGCTAAATTTTGTGATTTTAACTA	1143
Db	8642	CGAGTAGCTGGGATTA CAGGCGTGGCCAC CAGCTTGGCTAAATTTTGTGATTTTAAATA	8701
Qy	1144	GAGACGTGGTTCACATATGTTGGCCAGGCTGATCTCGAACTCCTGACCTTGTATCCACC	1203
Db	8702	GAGATGGGGTTTCAACCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTTGTGATCTGCC	8761
Qy	1204	CGCCTCGGCCTGCCAAAGTGCTGGGATTA CAGGCGTGAGCCATTTGTGCCTGGCCGA	1259
Db	8762	CGTCTTGGCCTCCCAAAGTGCTGGGATTA CAGGCGTGAGCCATTCGCGCTGGCCAA	8817

RESULT 14
US-09-780-172-18/c
; Sequence 18, Application US/09780172
; Patent NO. 6607916
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
; FILE REFERENCE: RTS-0159
; CURRENT APPLICATION NUMBER: US/09/780,172
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 63000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-172-18

Query Match	8.9%	Score	226.6	DB 4	Length	63000			
Best Local Similarity	81.6%	Pred. No.	7.8e-35						
Matches	262	Conservative	0	Mismatches	59	Indels	0	Gaps	0
Qy	961	AGATTTTTTTGTTGTTTTTGACCGAGGTCTACTCTGTCAACCAGGCTGGAGTGAATCGC	1020						
Db	43854	AGCCTTTTTTTTTTTTGACACAGGCTTGCTCTGTGCCAGGCTGGAGTGTAGTACG	43795						
Qy	1021	ATGTCCTCAGCTCACTGCAACCTCCGCTCCGGGTTCAGGCCATCTCTTCACCTCAGCC	1080						
Db	43794	ATGATCTCAGCTCACTGCAACTTCGGCTCCGGGTTCAGGCAATCTCTCGCTCAGCC	43735						
Qy	1081	TCCCAAGTAGCTGGACTACAGGCATGAGCCACCGCACCCGGGTAAATTTTGTATTTTTA	1140						
Db	43734	TCCCGAGTAGCTGGGATTACAGCGGCCAGCCACATGCCCCAGCTAAATTTTGTATTTTTA	43675						
Qy	1141	GTAGAGACGTGGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCTGTAAATCC	1200						
Db	43674	GTAGAGACAGGGTTTTCATCATGTGTGTGACGATCGTCTCGAACTCTCGACCTCATGATCC	43615						
Qy	1201	ACCCGCCTTCGGCCTCCAAAAGTGTGGGATTACAGCGGTGAGCCATTTGTGCTGGCCGAT	1260						
Db	43614	GCCTGCCTTGGCCTTCCAAAGTGTGGGATTATAGGCATGAGCCACTGTGCTGGCCTTA	43555						
Qy	1261	TTTTTAAAAAATGATTTCTTA	1281						
Db	43554	TATTATAAAGCCTTCTCTCGA	43534						

RESULT 15
US-09-816-093-3/c
; Sequence 3, Application US/09816093
; Patent No. 6518055
; GENERAL INFORMATION:
; APPLICANT: GAN, weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

```

, TITLE OF INVENTION:  USES THEREOF
,
, FILE REFERENCE:  CLO01182
,
, CURRENT APPLICATION NUMBER:  US/09/816,093
,
, CURRENT FILING DATE:  2001-03-26
,
, NUMBER OF SEQ ID NOS:  4
,
, SOFTWARE:  FastSEQ for Windows Version 4.0
,
, SEQ ID NO 3
,
, LENGTH:  46718
,
, TYPE:  DNA
,
, ORGANISM:  Human
,
, FEATURE:
,
, NAME/KEY:  misc_feature
,
, LOCATION:  (1)...(46718)
,
, OTHER INFORMATION:  n = A,T,C or G
,
, US-09-816-093-3
,

```

Query Match	8.9%;	Score 226.2;	DB 4;	Length 46718;	
Best Local Similarity	67.2%;	Prod. No. 8.9e-35;			
Matches 354;	Conservative	0;	Mismatches 163;	Indels 10;	Gaps 2;
QY	960	TAGATTTTTTTGTGTTTTTTGACACGAGCTTACTCTGTACCCAGGCTGGAGTGAATGG	1019		
DB	15830	TTGTCCTTTTTTTTGTGTTTTTTGAGCGGAGCTTCGCTCTGT-TGCAGGCTGGAGTTCAGTGG	15772		
QY	1020	CATGCTCTCAGCTCACTGCACCTCCGCTCCCGGTTCAAGCCATTCCTCCACCTCAGC	1079		
DB	15771	CGCGTCTCGGCTCACTGCACCTCCGCTCCGCTCTGGTTCAGAGTATCTCTCGCTCAGC	15712		
QY	1080	CTCCCAAGTAGCTGGGACTACAGGATGAGCACCGCACCGGCTAAATTTTGTATTTTTT	1139		
DB	15711	CTCCCAAGTAGCTGGGATTTACAGGCACACGCCACCATGCTGGCTAAATTTTGTATTTTTT	15652		
QY	1140	AGTAGAGCGTGGTTCCCATATGTTGGCCAGGCTGATCTCGAACTCCTGACCTTTGTAATC	1199		
DB	15651	AGTAGAGATGGGGTTTCACTATGTTGGCCAGGCTGGCTCTTGAACCTTCTGACCTTTGTATC	15592		
QY	1200	CACCGGCTCTCGGCTTGCACCAAGTGTGGGATTTACAGCGCTGAGCCATTTGTCCCTGGCCG-	1258		
DB	15591	CACCTCTCTCGGCTTCCCAAGTGTGGGATTTACAGGCATTAAGCCACTGTGCCTGGCCGT	15532		
QY	1259	-----ATTTTTTAAAAATGTAATCTTATGTCTAGTTTTTCAATAGTTTTTATTTAAAT	1310		
DB	15531	ATTTAGGTATTTTTTTTAAAGCTCTCTTTCTGTTACTGGATTTTTTATCTTTATTTCTTGAC	15472		
QY	1311	GCAITTTTCCATTTGATGTAAGCTTTCAAAATTTATAGTATAGTTGTTCCTAGTATTTCTTT	1370		
DB	15471	ATTGAGAGAATCTAACCTCAGAGTTACACATTAGGAAAATTTATTAAAGATAGTTTTTCAT	15412		
QY	1371	ATCTTTTGTAACTCTGTTCAGCGCTCTGTAAGTGGCCTCTTTTTTAATAAATAATATTATTT	1430		
DB	15411	ACCTTAGGTAAACACAGCTCAGTTTATTACATCTTCCAACATAGCTAAACGATCA	15352		
QY	1431	GTTTCGCGTTTTTGTCTATTTTTTTTCTTATGCTCTTTGAGAGGATAT	1477		
DB	15351	GCTGCTCTCTCTGCTCTCTGCAATAACATTTCTTTTAGAGGGTTAT	15305		

Search completed: May 21, 2004, 23:07:13
Job time : 202.141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:24:44 ; Search time 1052.27 Seconds
(without alignments)
10258.453 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcataattttctttaa.....cttctctagtgtgaatttt 2541

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	335	13.2	335	3	Aa94789 Human Pro
2	276	10.9	59215	7	Aad36834 Human tra
3	271.2	10.7	174566	7	Abq77400 Human ITG
4	255.8	10.1	393	4	Aa181268 Human pol
5	252	9.9	157875	6	ABK99972 Human
6	244.2	9.6	9573	9	ADb31452 Bicalutam
7	241.4	9.5	110000	9	ADb31452 Bicalutam
8	239.8	9.4	21470	8	Abk42270 Genomic S
9	239.8	9.4	21470	8	ADb60426 Connectiv
10	239.4	9.4	16086	4	AaK69721 Human imm
11	239.4	9.4	16086	4	AaK69721 Human imm
12	239.4	9.4	16086	7	Abx59008 cDNA enco
13	238.8	9.4	538	4	Aa104865 Human rep
14	238.8	9.4	538	4	Aa104867 Human rep
15	238.8	9.4	538	4	Ab197761 Human tes
16	238.8	9.4	538	4	Ab197759 Human tes
17	238.4	9.4	10587	3	AaK92382 Human imm
18	238.2	9.4	43795	3	AaZ92583 Human DAZ
19	238.2	9.4	83391	6	Abq67094 Human ang
20	238	9.4	12729	5	Aaf97873 Human neu
21	238	9.4	15266	4	AaK73549 Human imm
22	238	9.4	15266	4	Aa137432 Human mus
23	238	9.4	15266	7	Abx60420 cDNA enco

C 24	238	9.4	15271	4	AAK73550 Human imm
C 25	238	9.4	15271	4	Aa137433 Human mus
C 26	238	9.4	15271	7	Abx60421 cDNA enco
C 27	238	9.4	143068	3	AAA34983 Human ade
C 28	238	9.4	143068	3	AAA35150 Human low
C 29	238	9.4	143068	3	Aaf21272 Human low
C 30	238	9.4	143068	3	Aaf21105 Human low
C 31	238	9.4	143068	6	Ab168124 Ovary can
C 32	238	9.4	143068	6	Abt11034 Human bre
C 33	238	9.4	143068	7	Abz96966 Human nuc
C 34	238	9.4	143068	7	Abz96799 Human nuc
C 35	238	9.4	149412	3	AAA35151 Human ade
C 36	238	9.4	152740	3	Aaf21273 Human low
C 37	238	9.4	152740	7	Abz96967 Human nuc
C 38	237.8	9.4	126512	6	Abn83429 Human tra
C 39	237.2	9.3	31952	4	AaK89370 Human dig
C 40	236.4	9.3	62804	6	AAD39317 Human cal
C 41	236.4	9.3	62804	7	Abx10916 Genomic D
C 42	235.6	9.3	17590	3	AaZ50904 Human TBC
C 43	235.4	9.3	185035	6	Abt10147 Human bre
C 44	235.4	9.3	185035	7	Aca64951 Human PEN
C 45	234.8	9.2	13444	4	Aa137295 Human mus

ALIGNMENTS

RESULT 1

AAA94789
ID AAA94789 standard; DNA; 335 BP.

XX
AC AAA94789;

XX
DT 19-JAN-2001 (first entry)

XX
DE Human Prostate-Specific Membrane antigen Enhancer PSME core region.

XX
KW Core region; human; prostate-specific membrane antigen; enhancer; PSME;
KW cancer; renal cell carcinoma; transitional cell; colonic; neuroendocrine;
KW malignant melanoma; pancreatic duct; breast; soft tissue; PSMA;
KW non-small cell lung; testicular embryonal; glioblastoma multiforme;
KW prostate; breast; bladder; cancer; ss.

XX
OS Homo sapiens.

XX
PN W0200052156-A1.

XX
PD 08-SEP-2000.

XX
PF 01-MAR-2000; 2000WO-AU000143.

XX
PR 01-MAR-1999; 99AU-00008956.

XX
PR 25-JAN-2000; 2000AU-00005268.

XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX
PI Molloy PL, Watt F;

XX
DR WPI; 2000-594182/56.

XX
PT New regulatory constructs comprising intron 3 of the prostate-specific
PT membrane antigen gene and a heterologous peptide, for directing gene
PT expression in a prostate, bladder, breast or vascular endothelial cell.

XX
PS Claim 10; Fig 11; 56pp; English.

XX
CC The present sequence is the core region from human prostate-specific
CC membrane antigen enhancer (PSME). This sequence is located in intron 3 of
CC the PSMA gene. This sequence was used to construct a recombinant
CC expression cassette, which is useful for the expression of proteins in
CC vascular endothelial cells. In addition, the recombinant construct is
CC useful in the treatment of cancer e.g. renal cell carcinoma, transitional
CC cell carcinoma, colonic adenocarcinoma, neuroendocrine carcinoma,

CC malignant melanoma, pancreatic duct carcinoma, breast carcinoma, soft
CC tissue carcinoma, non-small cell lung carcinoma, testicular embryonal
CC carcinoma, glioblastoma multiforme, prostate, breast or bladder cancer
XX
SQ Sequence 335 BP; 84 A; 55 C; 54 G; 142 T; 0 U; 0 Other;
Query Match 13.2%; Score 335; DB 3; Length 335;
Best Local Similarity 100.0%; Pred. No. 4.6e-40;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 721 AATTATTTTCTCTTAACCTTCAAACTCAAGGAAACAGTGGCTTGACTCTGTTT 780
DB 1 AATTATTTTCTCTTAACCTTCAAACTCAAGGAAACAGTGGCTTGACTCTGTTT 60
QY 781 GTGGAATTTTAAACTACTGTTTAAATTTCTTTATTTGTTGTAATATGACTATTTTACG 840
DB 61 GTGGAATTTTAAACTACTGTTTAAATTTCTTTATTTGTTGTAATATGACTATTTTACG 120
QY 841 TCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAAATTTTA 900
DB 121 TCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAAATTTTA 180
QY 901 TGTTCATAGAACATTTTATTTGTTGTCATATGACTTTATTTGTTGTCATATGATAAATTTTA 960
DB 181 TGTTCATAGAACATTTTATTTGTTGTCATATGACTTTATTTGTTGTCATATGATAAATTTTA 240
QY 961 AGATTTTGTGTTTGTGTTTGAACGAGTCTTACTCTGTACCCAGGCTGGAGTGAATGGC 1020
DB 241 AGATTTTGTGTTTGTGTTTGAACGAGTCTTACTCTGTACCCAGGCTGGAGTGAATGGC 300
QY 1021 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 1055
DB 301 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 335
RESULT 2
AAD36834/c
ID AAD36834 standard; DNA; 59215 BP.
XX AC AAD36834;
XX DT 07-MAR-2003 (first entry)
XX DE Human transporter protein genomic DNA.
XX KW Human; transporter; sodium-glucose cotransporter; chromosome 16;
XX KW drug screening; single nucleotide polymorphism; SNP; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 3000..57168
XX FT /tag= a
XX FT /product= "Transporter protein"
XX FT /note= "The CDS is interrupted by 12 introns"
XX FT 3000..3014
XX FT /tag= b
XX FT /label= 1
XX FT 3015..10042
XX FT /tag= c
XX FT /label= 1
XX FT /replace(4429,A)
XX FT /tag= aa
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /replace(5167,A)
XX FT /tag= ab
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /replace(6198,C)
XX FT /tag= ac
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /replace(6760,A)
XX FT /tag= ad
XX FT /standard_name= "Single nucleotide polymorphism"

FT variation replace(6934, A/-)
FT /tag= ae
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "The nucleotide is replaced by A or no nucleotide"
FT variation replace(7038,C)
FT /tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(7116,T)
FT /tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(7978,G)
FT /tag= ah
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(9973,A)
FT /tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT 10043..10147
FT /tag= d
FT /label= 2
FT 10148..12305
FT /tag= e
FT /label= 2
FT variation replace(11318,T)
FT /tag= aj
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(11998,T)
FT /tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(12047,T)
FT /tag= al
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(12058,T)
FT /tag= am
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(12141,G)
FT /tag= an
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(12163,T)
FT /tag= ao
FT /standard_name= "Single nucleotide polymorphism"
FT 12306..12365
FT /tag= f
FT /label= 3
FT 12366..12408
FT /tag= g
FT /label= 3
FT variation replace(19275,G)
FT /tag= ap
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(2110,A)
FT /tag= ar
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(21118,C)
FT /tag= as
FT /standard_name= "Single nucleotide polymorphism"
FT 21409..21513
FT /tag= h
FT /label= 4
FT 21514..23038
FT /tag= i
FT /label= 4
FT /replace(22575,A)
FT /tag= at
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(22914,T)
FT /tag= au
FT /standard_name= "Single nucleotide polymorphism"
FT 23039..23144
FT /tag= j
FT /label= 5
FT /replace(23106,C)
FT /tag= av
FT /standard_name= "Single nucleotide polymorphism"

[illegible]

RESULT 7			
ADE11169_2			
Continuation (3 of 4) of ADE11169 from base 200001 (Human transporter protein encoding gene)			
WP	Sequence split into 4 fragments	LOCUS ADE11169	Accession ADE11169
WP	Fragment Name	Begin	End
WP	ADE11169_0	1	110000
WP	ADE11169_1	100001	210000
WP	ADE11169_2	200001	310000
WP	ADE11169_3	300001	394191

[illegible]

RESULT 8
ABK42270
ID ABK
XX
AC ABK
XX
DT 21-

XX	Genomic sequence #169 encoding novel human connective tissue polypeptide
DE	
XX	
XX	Human; connective tissue related disorder; cancer; gene therapy;
KW	cytostatic; gene; ds.
KW	
XX	
XX	Homo sapiens.
OS	
XX	
PN	WO200155343-A1.
XX	
XX	
PD	02-AUG-2001.
PD	
XX	
PP	17-JAN-2001; 2001WO-US001322.

PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234937P.
 PR 25-SEP-2000; 2000US-0234938P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 03-DEC-2000; 2000US-0251030P.
 PR 03-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 24533; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AM82170 to AM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially

PR 13-OCT-2000; 2000US-02399335P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
XX cancers and also for testing and detection e.g. diagnosis.
XX
XX Example 2; SEQ ID NO 2385; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins

CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, multiple sclerosis, rheumatoid
CC diabetes mellitus, Crohn's disease, (c) cardiovascular disorders such as
CC arthritis and ulcerative colitis; (d) wound healing; (e) neurological diseases e.g.
CC myocardial ischaemias; (f) and (g) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 16086 BP; 3365 A; 4404 C; 4642 G; 3675 T; 0 U; 0 Other;

Query Match 9.4%; Score 239.4; DB 4; Length 16086;
Best Local Similarity 70.8%; Pred. No. 2.5e-26;
Matches 318; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
QY 817 TGGTTGTAATACACTATTTTACGTCAATATAACAATTTTATTTGTTGTTAAATGACTTT 876
DB 13032 TGTAAATGTTTGAAGTGTGATAGGCAATTTAGTCTCAATGCTGATTTTATATATGAT 13091
QY 877 ATGTTTGTCAATGATAATTTTATGTCATAGAACAAATTTTATGCTTGATATATGACT 936
DB 13092 ATATTGGGTTATATCTCCCATCACATTTTGTGCTATTGTCACCTCTGTGTTCT 13151
QY 937 TTATGTTTATAGGCTATACAACTAGATTTTTTTGTTGTTTGTGACCGAGTCTTACTCTG 996
DB 13152 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 13211
QY 997 TCACCCAGGCTGAGTGTATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
DB 13212 TCACCCAGGCTGAGTGTGCAATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13271
QY 1057 TCAAGCCATTCTTCCACCTCAGCCTCCCAAGTAGCTGGGACTACAGGCATGACCCACCG 1116
DB 13272 TCATGCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13331
QY 1117 ACCCGGCTAATTTTGTATTTTATAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCT 1176
DB 13332 GGGCGGCTAATTTTGTATTTTATAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCT 13391
QY 1177 CTGAACTCTGACCTGTGTAATCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1236
DB 13392 CTCGAACTCTGACCTGTGTAATCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13451
QY 1237 CGTGAGCCATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265
DB 13452 CGTGAGCCATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13480

RESULT 12
ABX59008
ID ABX59008 standard; cdna; 16086 BP.
XX
XX AC ABX59008;
XX
XX DT 26-FEB-2003 (first entry)
XX
XX DE cDNA encoding novel human musculoskeletal system antigen #1352.
XX
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX post-operative tissue repair; limb regeneration; neuronal growth;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX AIDS-related complex; chondrocyte growth; bone regeneration;

KW periodontal regeneration; tissue transport; bone graft; skin aging;
KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KW cell growth; organ transplant; cell differentiation; body height; weight;
KW hair colour; eye colour; skin; percentage of adipose tissue;
KW pigmentation; cosmetic surgery; metabolism; biorhythm; circadian rhythm;
KW depression; tendency for violence; pain; reproductive capability;
KW hormone level; endocrine level; appetite; libido; memory; stress;
KW storage capability; fat content; lipid content; protein content;
KW carbohydrate content; vitamin content; cofactor content;
KW nutritional component.

XX Homo sapiens.

XX OS

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

17-JAN-2001; 2001US-00764877.
31-JAN-2000; 2000US-0179065P.
04-FEB-2000; 2000US-0180628P.
28-JUN-2000; 2000US-0214886P.
07-JUL-2000; 2000US-0216647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
14-AUG-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
20-AUG-2000; 2000US-0226868P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
08-SEP-2000; 2000US-0231413P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
27-SEP-2000; 2000US-0235834P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0244617P.
17-NOV-2000; 2000US-0249299P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2003-128199/12.

XX Isolated nucleic acid molecules encoding musculoskeletal system

XX associated polypeptides, useful for detecting disorders, e.g. cancer.

XX Disclosure; SEQ ID NO 2385; 321pp; English.

XX The invention describes an isolated nucleic acid molecule comprising a

XX sequence encoding musculoskeletal system associated polypeptides useful
XX for detecting disorders, e.g., cancer or cancer metastases, in animals or
XX humans. The nucleic acid: stimulates re-vascularisation of ischaemic
XX tissues associated with conditions such as thrombosis, arteriosclerosis,
XX and other cardiovascular conditions; treats wounds due to injuries,
XX burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
XX and limb regeneration; stimulates neuronal growth; can treat and prevent
XX neuronal damage occurring in certain disorders or neurodegenerative
XX conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-
XX related complex; stimulates chondrocyte growth, thus they can be used to
XX enhance bone and periodontal regeneration and aid in tissue transports or
XX bone grafts; prevents skin aging due to sunburn by stimulating
XX keratinocyte growth; prevents hair loss, since FGF family members
XX activate hair-forming cells and promotes melanocyte growth; stimulates
XX growth and differentiation of hematopoietic cells and bone marrow cells
XX when used in combination with other cytokines; maintains organs before
XX transplantation or for supporting cell culture of primary tissues;
XX induces tissue of mesodermal origin to differentiate in early embryos;
XX increases or decreases the differentiation or proliferation of embryonic
XX stem cells, besides, haematopoietic lineage; modulates mammalian
XX characteristics, such as, body height, weight, hair colour, eye colour,
XX skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,
XX state or physical state by influencing metabolism; changes mammal's metal
XX depression, tendency for violence, tolerance for pain, libido, memory, or
XX capabilities, hormonal or endocrine levels, appetite, fat content, lipid,
XX stress; increases or decreases storage capabilities, cofactors or other nutritional
XX protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
XX components. This sequence encodes a novel human musculoskeletal system
XX antigen. Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?DocID=20020147140

XX Sequence 16086 BP; 3365 A; 4404 C; 4642 G; 3675 T; 0 U; 0 Other;

XX Query Match 9.4%; Score 239.4; DB 7; Length 16086;

XX Best Local Similarity 70.8%; Pred. No. 2.5e-26;

XX Matches 318; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

XX QY 817 TGGTTGTAATGACTATTTTACGTATATACAAATTTTATTTTGTGTTAAATGACTTT 876

XX Db 13032 TGTTAATGTTTGACTGTGTAGGCAATTTAGTTTCACTGATGCTGTATTTTATTATGAT 13091

XX QY 877 ATTCTTTGTCATATGATAATTTTATGTCATAGACAAATTTTATGCTGTATATGACT 936

XX Db 13092 ATATTGGTTTATATCCCAATTCACATTTGCTATTGTCCCACTCTGTGTTTCT 13151

XX QY 937 TTAATGTTATATGGCTATACAACTAGATTTTTTTTGTGTTTGTGACCACTTACTCTG 996

XX Db 13152 TTTTCTCTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 13211

XX QY 997 TCACCCAGGCTGGAGTGAATGGCATGGTCTCAGCTCAGTGCACACCTCCGCTCCCGGGT 1056

XX Db 13212 TCACCCAGGCTGGAGTGAATGGCATGGTCTCAGCTCAGTGCACACCTCCGCTCCCGGGT 13271

XX QY 1057 TCAAGCCATTCTCCACCTCAGCTCCAGTAGCTGGGACTACAGCATGAGCCACCGC 1116

XX Db 13272 TCATGCCATCTCTCTGCTCAGCTCCCAAGTAGCTGGGATTACAGGCGCTGCCACCAT 13331

XX QY 1117 ACCCGGCTAAATTTTGTATTTTAGTAGACCTGGTTCCACTATGTTGCCAGGCTGAT 1176

Db	13332	GGCGGCTAATATTTGTAATTTTATAGTAGACACGGGTTTACCATGTTGGTCAGGCTGGT	13391
Qy	1177	CTCGAACTCTTGACCTTGTAAATCCACCGCGCTCGGCTGCCAAAGTCTGGGATTACAGG	1236
Db	13392	CTCGAACTCTGACCTTGTGATCTGCCGCTTGGCTCCCAAAGTCTGGGATTACAGG	13451
Qy	1237	CGTGAGCCATGTGCTGGCCGATTTTTT	1265
Db	13452	CGTGAGCCACCGCGCTGGCTTTTCT	13480
RESULT 13			
AAL04865/c			
ID	AAL04865 standard; DNA; 538 BP.		
AC	AAL04865;		
DT	21-NOV-2001 (first entry)		
DE	Human reproductive system related antigen DNA SEQ ID NO: 7553.		
KW	Human; reproductive system related antigen; reproductive system disorder;		
KW	cancer; gene therapy; ds.		
OS	Homo sapiens.		
PN	WO200155320-A2.		
PD	02-AUG-2001.		
XX	17-JAN-2001; 2001WO-US001339.		
XX	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-01981123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225826P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225577P.		
PR	18-AUG-2000; 2000US-0225759P.		
PR	22-AUG-2000; 2000US-0226661P.		
PR	22-AUG-2000; 2000US-0226868P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		
PR	30-AUG-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR	05-SEP-2000; 2000US-0229509P.		
PR	05-SEP-2000; 2000US-0229510P.		
PR	05-SEP-2000; 2000US-0229511P.		
PR	05-SEP-2000; 2000US-0229512P.		
PR	05-SEP-2000; 2000US-0229513P.		
PR	05-SEP-2000; 2000US-0229514P.		
PR	05-SEP-2000; 2000US-0229515P.		
PR	05-SEP-2000; 2000US-0229516P.		
PR	05-SEP-2000; 2000US-0229517P.		
PR	05-SEP-2000; 2000US-0229518P.		
PR			

PR 17-NOV-2000; 2000US-02492117P.
PR 17-NOV-2000; 2000US-02492118P.
PR 17-NOV-2000; 2000US-02492444P.
PR 17-NOV-2000; 2000US-02492455P.
PR 17-NOV-2000; 2000US-02492649P.
PR 17-NOV-2000; 2000US-02492655P.
PR 17-NOV-2000; 2000US-02492977P.
PR 17-NOV-2000; 2000US-02492999P.
PR 17-NOV-2000; 2000US-02493000P.
PR 01-DEC-2000; 2000US-02501600P.
PR 01-DEC-2000; 2000US-02503919P.
PR 05-DEC-2000; 2000US-02510302P.
PR 05-DEC-2000; 2000US-02519889P.
PR 05-DEC-2000; 2000US-02567119P.
PR 06-DEC-2000; 2000US-02514799P.
PR 08-DEC-2000; 2000US-02518569P.
PR 08-DEC-2000; 2000US-02518699P.
PR 08-DEC-2000; 2000US-02518699P.
PR 08-DEC-2000; 2000US-02519899P.
PR 08-DEC-2000; 2000US-02519900P.
PR 11-DEC-2000; 2000US-02540979P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
XX
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 7553; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
XX Sequence 538 BP; 151 A; 134 C; 139 G; 114 T; 0 U; 0 Other;
SQ

Query Match 9.4%; Score 238.8; DB 4; Length 538;
Best Local Similarity 89.0%; Pred. No. 4e-26;
Matches 258; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 967 TTTTGTGTTGTTTTCACCGAGTCTTACTCTGTACCCAGGCTGGAGTGAATGGCATGGTC 1026
Db 533 TTTTGTGTTTTCACAGAGTCTACTCTGTACCCAGGCTGGAGTGAATGGCATGGTC 474

Qy 1027 TCAGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCCATTCTCCACCTCAGCCTCCCAA 1086
Db 473 TCCACTCACTGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCACCTCAGCCTCCCGA 414

Qy 1087 GTAGCTGGGACTACAGGATGAGCAGCCGCGGCTAATTTTGTATTTTATAGTAGAG 1146
Db 413 GTAGCTGGGACTACAGGATGAGCAGCCGCGGCTGATTTTGTATTTTATAGTAGAG 354

Qy 1147 ACGTGGTTCACACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCTTCTATCCACCGC 1206
Db 353 ATGGGGTTTCACTAATGTTGGCCAGGCTGCTCTAACTCTCTGACCTCATGATCCACCTGC 294

Qy 1207 CTCGGCTGCCAAAGTGTGGGATTACAGGCGTGAGCATTGTGCTGGC 1256
Db 293 CTCGGCTGCCAAAGTGTGGGATTACAGGCGTGAGCATTGTGCTGGC 244

RESULT 14
AAL04867/c
ID AAL04867 standard; DNA; 538 BP.
XX
AC AAL04867;
XX

DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 7555.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234224P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-023935P.
 PR 13-OCT-2000; 2000US-023937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249284P.
 PR 17-NOV-2000; 2000US-0249285P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465570/50.
 DR
 XX
 XX Isolated nucleic acid molecule encoding a reproductive system antigen is
 PT used in preventing, treating or ameliorating a medical condition.
 XX
 XX Disclosure; SEQ ID NO 7555; 1297pp + Sequence Listing; English.
 XX
 XX The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention
 XX
 XX Sequence 538 BP; 151 A; 134 C; 139 G; 114 T; 0 U; 0 Other;
 SQ
 Query Match 9.4%; Score 238.8; DB 4; Length 538;
 Best Local Similarity 89.0%; Pred. No. 4e-26;
 Matches 258; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 967 TTTTGTGTTTGGACCGAGTCTTACTCTCTACCCAGGCTGGAGTGAATGGCATGGTC 1026
 DB 533 TTTTTTTTTTTCAGACAGTCTCCTCTCTACCCAGGCTGGAGTGAATGGCATGGTC 474
 QY 1027 TCAGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCAATTTCCACCTCAGCTCCCAA 1086
 DB 473 TCCACTCACTGCAACCTCCGCTCCCGGGTTCAAGCAATTTCCACCTCAGCTCCCGA 414
 QY 1087 GTAGCTGGGACTACAGGCATGAGCCACCGCAGGCTAAATTTGTATTTTAGTAGAG 1146
 DB 413 GTAGCTGGGACTACAGGCATGAGCCACCGCAGGCTAAATTTGTATTTTAGTAGAG 354
 QY 1147 ACCTGGTTCACATATGTTGCCAGGCTGATCTCGAACTCTGACCTTGTATTCACCCGC 1206
 DB 353 ATGGGGTTTCACTATGTTGCCAGGCTGTTCTCAACTCTGACCTCATGATCCACTGC 294
 QY 1207 CTCGGCTGCCAAAGTCTGGGATTACAGCGGTGAGCCATTGTCCTGGC 1256
 DB 293 CTCGGCTGCCAAAGTCTGGGATTACAGCGGTGAGCCACCGCAGCTGGC 244
 RESULT 15
 ABL97761/c
 ID ABL97761 standard; DNA; 538 BP.
 XX
 AC ABL97761;
 XX
 DT 21-JUN-2002 (first entry)
 XX
 DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2413.
 XX
 KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disease; infection; cytostatic; gene; ds.
 OS Homo sapiens.
 XX
 XX W0200155317-A2.
 XX
 XX 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US001329.
 XX

PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198121P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217966P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 06-DEC-2000; 2000US-0251799P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:26:10 ; Search time 10196.7 Seconds
(without alignments)
10801.019 Million cell updates/sec

Title: US-09-914-651A-2_COPY_14040_16580

Perfect score: 2541

Sequence: 1 agcataattttcttttaatt.....cttctctagtgtgaatttt 2541

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2541	100.0	93525	9	AF007544	AF007544 Homo sapi
2	2510	98.8	129095	9	AC110742	AC110742 Homo sapi
3	2508.4	98.7	158524	2	AL162372	AL162372 Homo sapi
c	2508.4	98.7	187529	9	AC118273	AC118273 Homo sapi
5	2367	93.2	135637	9	AF003400	AF003400 Homo sapi
6	2367	93.2	137888	9	AP003122	AP003122 Homo sapi
c	2367	93.2	192648	2	AC024234	AC024234 Homo sapi
8	2148	84.5	117521	2	AC023784	AC023784 Homo sapi
c	1476.6	58.1	246865	2	AF074003	AF074003 Homo sapi
10	1461.4	57.5	1669	9	AF480875	AF480875 Homo sapi
11	930.8	36.6	246865	2	AC074003	AC074003 Homo sapi
c	653	25.7	68842	2	AC137087	AC137087 Homo sapi
12	603.8	23.8	72409	2	AC131257	AC131257 Homo sapi
c	346.2	13.6	40131	2	AC136711	AC136711 Homo sapi
14	335	13.2	335	6	BD269371	BD269371 Controili
15	321.2	12.6	175342	9	AC108696	AC108696 Homo sapi
16	321.2	12.6	175233	2	AC022941	AC022941 Homo sapi
17	320.2	12.6	49838	9	HS387E22	HS387E22 Human DNA
18	312	12.3	192802	9	AC005284	AC005284 Homo sapi
19	308.6	12.1	141742	9	AC092621	AC092621 Homo sapi
c	303.8	12.0	192104	9	AC006994	AC006994 Homo sapi
21	298	11.7	87097	9	AC109344	AC109344 Homo sapi
22	298	11.7	169261	2	AC140480	AC140480 Homo sapi
23	296.8	11.7	156957	9	AP001977	AP001977 Homo sapi
24	296.8	11.7	171752	2	AC147307	AC147307 Pan trogl
c	294.4	11.6	167796	2	AC109819	AC109819 Homo sapi
26	293.4	11.5	98842	9	AC008972	AC008972 Homo sapi
c	293.4	11.5	108162	2	AC020960	AC020960 Mus muscu
28	293.4	11.5	124235	9	AC020893	AC020893 Homo sapi
c	293.4	11.5	152905	2	AC044801	AC044801 Homo sapi
30	293.4	11.5	198027	2	AC008515	AC008515 Homo sapi
31	292.8	11.5	105211	9	AC110076	AC110076 Homo sapi
c	292.6	11.5	125402	9	AC114815	AC114815 Homo sapi
32	291.6	11.5	83248	9	AC068754	AC068754 Homo sapi
33	291.6	11.5	173724	2	AC026766	AC026766 Homo sapi
c	291	11.5	172705	9	AC013272	AC013272 Homo sapi
36	290.2	11.4	78794	9	AC003990	AC003990 Homo sapi
c	287.4	11.3	41572	9	AY217547	AY217547 Homo sapi
38	287.4	11.3	159446	9	AC003669	AC003669 Homo sapi
39	287.4	11.3	190752	2	AC114805	AC114805 Homo sapi
40	287.4	11.3	194492	9	AC097625	AC097625 Homo sapi
41	286	11.3	109601	9	AC074392	AC074392 Homo sapi
42	284.6	11.2	170940	9	AC120042	AC120042 Homo sapi
43	284.6	11.2	181694	9	AC019357	AC019357 Homo sapi
44	282.2	11.1	187883	9	AC008268	AC008268 Homo sapi
c	282.2	11.1	187883	9	AC008268	AC008268 Homo sapi

ALIGNMENTS

RESULT 1
AF007544
LOCUS
DEFINITION Homo sapiens prostate-specific membrane antigen (PSM) gene,
complete cds.
ACCESSION AF007544
VERSION AF007544.1 GI:2970122
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 93525)
O'Keefe,D.S., Su,S.L., Bacich,D.J., Horiguchi,Y., Luo,Y.,
Powell,C.T., Zandvliet,D., Russell,P.J., Molloy,P.L., Nowak,N.J.,
93525 bp DNA linear PRI 10-DEC-1998

[illegible]

```
QY 2040 TTTAATATTTTGTATTTGACATTAGTTTGGCTTCTTATTTAGTTTGGCTGAAATCTCCA 2099
|
|
|
Db 124690 TTTAATATTTTGTATTTGACATTAGTTTGGCTTCTTATTTAGTTTGGCTGAAATCTCCA 124749
|
|
|
QY 2100 TTATCCTTAATGTGCTCTCTCATTTTGTCTGCTTCTTATTTAATTTAGAGATAAATGTTA 2159
|
|
|
Db 124750 TTATCCTTAATGTGCTCTCTCATTTTGTCTGCTTCTTATTTAATTTAGAGATAAATGTTA 124809
|
|
|
QY 2160 AATTATCTCAGCTCAGTATAGTATGCTCTGTTTATATATATATATATAAATTTATAAT 2219
|
|
|
Db 124810 AATTATCTCAGCTCAGTATAGTATGCTCTGTTTATATATATATATATAAATTTATAAT 124869
|
|
|
QY 2220 CCATRAAATTTATGTATATATTTTGGAGACCTTATATCATATATAAAGAGAAATGTTG 2279
|
|
|
Db 124870 CCATRAAATTTATGTATATATTTTGGAGACCTTATATCATATATAAAGAGAAATGTTG 124929
|
|
|
QY 2280 ATGAATGACAGACTTATATCTATGATAGTATGCTTATTTTATCTCGTCATAATGTTATTG 2339
|
|
|
Db 124930 ATGAATGACAGACTTATATCTATGATAGTATGCTTATTTTATCTCGTCATAATGTTATTG 124989
|
|
|
QY 2340 ACTTGTCTCTAAATTTTATTAATTAATTTTGTGTTTCTTTTTCAGCGGGTTT 2399
|
|
|
Db 124990 ACTTGTCTCTAAATTTTATTAATTAATTTTGTGTTTCTTTTTCAGCGGGTTT 125049
|
|
|
QY 2400 ATGTACTCTGCTGCTCAATTTGTTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAA 2459
|
|
|
Db 125050 ATGTACTCTGCTGCTCAATTTGTTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAA 125109
|
|
|
QY 2460 TTATCTCTTTTCCATTTTCAATTTTATTAATTTCTGATATACATAATTTAGTCACTTTT 2519
|
|
|
Db 125110 TTATCTCTTTTCCATTTTCAATTTTATTAATTTCTGATATACATAATTTAGTCACTTTT 125169
|
|
|
QY 2520 ACCTTCCTCTAGTGTGAATTTT 2541
|
|
|
Db 125170 ACCTTCCTCTAGTGTGAATTTT 125191
|
|
|
RESULT 3
LOCUS AL162372 158524 bp DNA linear HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome 13 clone RP11-181P16, 3 unordered pieces.
ACCESSION AL162372
VERSION AL162372.11 GI:12539555
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Burton, J.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jan 25, 2001 this sequence version replaced gi:11693478.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA181P16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 157639 bases at least Q40
Consensus quality: 157977 bases at least Q30
Consensus quality: 158153 bases at least Q20
Insert size: 158324; sum-of-contigs
Insert size: 174362; 5.7% error; agarose-fp
Quality coverage: 6.90x in Q20 bases; agarose-fp
coverage: 6.26x in Q20 bases; agarose-fp

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 59356: contig of 59356 bp in length
* 59357 59456: gap of 100 bp
* 59457 65248: contig of 5792 bp in length
* 65249 65348: gap of 100 bp
* 65349 158524: contig of 93176 bp in length.
FEATURES
Location/Qualifiers
1..158524
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="13"
/clone="RP11-181P16"
/clone_lib="RP11-11.1"
1..59356
/notes="assembly_fragment:00306
clone_end:17
vector_side:left"
59457..65248
/notes="assembly_fragment:01282
fragment_chain:1"
65349..158524
/notes="assembly_fragment:02815
fragment_chain:1"
ORIGIN
Query Match 98.7%; Score 2508.4; DB 2; Length 158524;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2534; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 1 AGCATAAATTTTCTTTTAACTTTACCTAGTAGTTCAGATGAGTTCAATTTCTCATGTGAAT 60
Db 70774 AGCATAAATTTTCTTTTAACTTTTACCTAGTAGTTCAGATGAGTTCAATTTCTCATGTGAAT 70833
QY 61 CACAGAAAAATATGCGGAAATTTTGAATATGTTGGGATGTGCAAGTGTGCAAGTGTGATTTT 120
Db 70834 CACAGAAAAATATGCGGAAATTTTGAATATGTTGGGATGTGCGGATGTGCGAAGTGTGATTTT 70893
QY 121 GGTTCATTTTACCTTCTATCCATGTCAGAAAAAGGAAATTAATCTGCTCTAGTCAGTA 180
Db 70894 GGTTCATTTTACCTTCTATCCATGTCAGAAAAAGGAAATTAATCTGCTCTAGTCAGTA 70953
QY 181 CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCAATTTGTATTAATCTG 240
Db 70954 CAATGTAAATCTGCCAGATCCCTCTCATCTCCAAATCCCTCCCAATTTGTATTAATCTG 71013
QY 241 CACTGACCAAGCCTATACAAATATATGTTGAAGAGAGTTTGTATGACAGATATCCTTA 300
Db 71014 CACTGACCAAGCCTATACAAATATATGTTGAAGAGAGTTTGTATGACAGATATCCTTA 71073
QY 301 ACTTTCCTTCTGGTCTTGAAGAAATACCTTTAAATCTTCTCCACCAATTAATTTGTTAT 360
Db 71074 ACTTTCCTTCTGGTCTTGAAGAAATACCTTTAAATCTTCTCCACCAATTAATTTGTTAT 71133
QY 361 TACTGTTTCTTTTCTGGTTATCTTAAAGTACATCTTTTCCACATATACATAAATTTT 420
Db 71134 TACTGTTTCTTTTCTGGTTATCTTAAAGTACATCTTTTCCACATATACATAAATTTT 71193
QY 421 TGCTGTATGCAATTTATTAACACATTTTCTTTTAAGATGAGCAGTCGATTTTCTTTC 480
Db 71194 TGCTGTATGCAATTTATTAACACATTTTCTTTTAAGATGAGCAGTCGATTTTCTTTC 71253
QY 481 TTGATCTGAAGATGACATTCAGTGATATTTCTAATGTTAGAAATATCTTATTTTGTGAA 540
Db 71254 TTGATCTGAAGATGACATTCAGTG- TATTTCTAATGTTAGAAATATCTTATTTTGTGAA 71312
```


ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 187529)

REFERENCE

Birren, B., Nussbaum, C. and Lander, E.

TITLE

Homo sapiens chromosome 11, clone RP11-107P7

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 187529)

AUTHORS

Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Landers, T., Lehotsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL

3 (bases 1 to 187529)

REFERENCE

Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zemdek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL

4 (bases 1 to 187529)

REFERENCE

Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zemdek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
Submitted (09-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2002 this sequence version replaced gi:22539138.
All repeats were identified using RepeatMasker:

COMMENT

Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26235
Center clone name: 107_P_7

FEATURES

source	Location/Qualifiers
1. .187529	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11"
	/clone="RP11-107P7"
	/clone_lib="RP11-107P7"
90. .113	/clone="RP11-107P7"
	/rpt_family="AT_rich"
1137. .1285	/rpt_family="MLT1E3"
	/rpt_family="MLT1E3"
1445. .1561	/rpt_family="MLT2F"
	/rpt_family="MLT2F"
2069. .2350	/rpt_family="AluJo"
	/rpt_family="AluJo"
3061. .3100	/rpt_family="(TATATG)n"
	/complement(3537..3585)
	/rpt_family="MLT2B1"
3608. .3880	/rpt_family="MER7A"
	/rpt_family="MER7A"
3881. .4194	/rpt_family="AluSx"
	/rpt_family="AluSx"
4195. .4245	/rpt_family="MER7A"
	/complement(4258..4386)
	/rpt_family="MLT2B3"
	/complement(4385..4623)
	/rpt_family="MLT2B3"
	/complement(4859..4953)
	/rpt_family="L1FA9"
	/complement(7027..7223)
	/rpt_family="L2"
7260. .7342	/rpt_family="(TATG)n"
	/complement(7349..7541)
	/rpt_family="L2"
	/complement(9158..9165)
	/note="<30 qual SNGL region"
	/complement(9182..9491)
	/rpt_family="AluY"
12672. .12990	/rpt_family="L1MB7"
	/rpt_family="L1MB7"
13060. .13216	/rpt_family="L1MB7"
	/rpt_family="L1MB7"
13402. .13694	/rpt_family="AluJo"
	/rpt_family="AluJo"
13750. .13778	/rpt_family="AT_rich"
	/complement(14039..20161)
	/rpt_family="L1PA5"
	/complement(21564..21843)
	/rpt_family="AluSp"
	/complement(23292..23499)

repeat_region	/rpt family="MER5A"	24668..24701	/rpt family="MER5A"	301	ACTTCTCTCTGCTCTTGAAGAAATACCTTTAAATCTCTTCTCACCATTAAATATGGTATT	360
repeat_region	/rpt family="TTTA)n"	complement(24705..24815)	/rpt family="TTTA)n"	Db		
repeat_region	/rpt family="AluSg/x"	27044..27313	/rpt family="AluSg/x"	170654	ACTTCTCTCTGCTCTTGAAGAAATACCTTTAAATCTCTTCTCACCATTAAATATGGTATT	170595
repeat_region	/rpt family="L1MB7"	33141..33378	/rpt family="L1MB7"	Qy	TACTGTTTCTTCTCTGCTCTTGAAGAAATACCTTTAAATCTCTTCTCACCATTAAATATGGTATT	420
repeat_region	/rpt family="MER2"	33379..33679	/rpt family="MER2"	Db	TACTGTTTCTTCTCTGCTCTTGAAGAAATACCTTTAAATCTCTTCTCACCATTAAATATGGTATT	170535
repeat_region	/rpt family="AluSc"	33680..33775	/rpt family="AluSc"	Qy	TGCTGATGATGATTTATATTAACACATTTTCTTTAAGATGAGCATGCGATTTCTTCTTC	480
repeat_region	/rpt family="MER2"	34199..34233	/rpt family="MER2"	Db	TGCTGATGATGATTTATATTAACACATTTTCTTTAAGATGAGCATGCGATTTCTTCTTC	170535
repeat_region	/rpt family="AT rich"	complement(34522..34554)	/rpt family="AT rich"	Qy	TTGATCTGAAGATGACATTCAGTGTATTTCTAATGTTAGAAATCTTATATTTTGA	540
repeat_region	/rpt family="AluSx"	34555..34576	/rpt family="AluSx"	Db	TTGATCTGAAGATGACATTCAGTGTATTTCTAATGTTAGAAATCTTATATTTTGA	170416
repeat_region	/rpt family="(TTG)n"	complement(34577..34857)	/rpt family="(TTG)n"	Qy	CTTGTAATATAGCATTAGGGTTTAAATTCATATATGATATATATATGTTTATACCTC	600
repeat_region	/rpt family="AluSg"	35428..35731	/rpt family="AluSg"	Db	CTTGTAATATAGCATTAGGGTTTAAATTCATATATGATATATATATGTTTATACCTC	170356
repeat_region	/rpt family="AluSg"	complement(35779..36074)	/rpt family="AluSg"	Qy	ATAGCATGTTTAGTCTGCTGTTTTCGAATCTATTTTGGGTATAATTTCAATATCTTA	660
repeat_region	/rpt family="AluSx"	36478..36530	/rpt family="AluSx"	Db	ATAGCATGTTTAGTCTGCTGTTTTCGAATCTATTTTGGGTATAATTTCAATATCTTA	170296
repeat_region	/rpt family="AT rich"	complement(36535..36767)	/rpt family="AT rich"	Qy	CTTGCTCTCTAAATGAGTTGGGTTTAAATTTCTGAAGTAGTCTTTATTTGCAATTA	720
repeat_region	/rpt family="AluSx"	36767..36767	/rpt family="AluSx"	Db	CTTGCTCTCTAAATGAGTTGGGTTTAAATTTCTGAAGTAGTCTTTATTTGCAATTA	170236
repeat_region	/rpt family="AluSx"	complement(37315..37760)	/rpt family="AluSx"	Qy	AATTATTTTCTCTTAACTCTGTTTAAATTTCTGAAGTAGTCTTTATTTGCAATTA	780
repeat_region	/rpt family="Tigger2a"	complement(38323..38470)	/rpt family="Tigger2a"	Db	AATTATTTTCTCTTAACTCTGTTTAAATTTCTGAAGTAGTCTTTATTTGCAATTA	170235
repeat_region	/rpt family="L2"	38475..38761	/rpt family="L2"	Qy	GTGGAATATTTTAACTCTGTTTAAATTTCTGAAGTAGTCTTTATTTGCAATTA	840
repeat_region	/rpt family="AluJb"	complement(38765..38850)	/rpt family="AluJb"	Db	GTGGAATATTTTAACTCTGTTTAAATTTCTGAAGTAGTCTTTATTTGCAATTA	170176
repeat_region	/rpt family="L2"	complement(38851..39161)	/rpt family="L2"	Qy	TCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA	900
repeat_region	/rpt family="AluY"	complement(39162..39261)	/rpt family="AluY"	Db	TCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA	170115
repeat_region	/rpt family="L2"	complement(40527..40838)	/rpt family="L2"	Qy	TGTCATAGAAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA	960
repeat_region	/rpt family="AluSx"	41030..41085	/rpt family="AluSx"	Db	TGTCATAGAAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA	170055
repeat_region	/rpt family="(TG)n"	complement(43618..43766)	/rpt family="(TG)n"	Qy	AGATTTTCTGTTTCTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA	169996
repeat_region	/rpt family="Tigger3 (Golem)"	43947..44091	/rpt family="Tigger3 (Golem)"	Db	AGATTTTCTGTTTCTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA	1020
repeat_region				Qy	ATGCTCTCAGCTCACTGCAACCTCCGCTCCGGGTTCAAGCCATTTCTCCACCTCAGCC	169936
repeat_region				Db	ATGCTCTCAGCTCACTGCAACCTCCGCTCCGGGTTCAAGCCATTTCTCCACCTCAGCC	1080
repeat_region				Qy	TCCCAAGTAGCTGGGATCTACAGGCATGAGCCACCGCCGCTAAATTTTGTATTTTA	169876
repeat_region				Db	TCCCAAGTAGCTGGGATCTACAGGCATGAGCCACCGCCGCTAAATTTTGTATTTTA	1140
repeat_region				Qy	GTAGAGAGCTGGTTCACACTATGTTGGCCAGGCTGATCTCGAACTCTCGACCTGTAATCC	169816
repeat_region				Db	GTAGAGAGCTGGTTCACACTATGTTGGCCAGGCTGATCTCGAACTCTCGACCTGTAATCC	1200
repeat_region				Qy	ACCGCTCCGGCTCCCAAGTAGCTGGGATTTACAGGCGTGGGATTTGCTGCGCCGAT	169756
repeat_region				Db	ACCGCTCCGGCTCCCAAGTAGCTGGGATTTACAGGCGTGGGATTTGCTGCGCCGAT	1260
repeat_region				Qy	TTTTTAAAAAATGATTTCTTATGTCAGTTTCTAATGTTTATTTAAAAATGCAATTTTCCA	169696
repeat_region				Db	TTTTTAAAAAATGATTTCTTATGTCAGTTTCTAATGTTTATTTAAAAATGCAATTTTCCA	1320
repeat_region				Qy	TTTTTAAAAAATGATTTCTTATGTCAGTTTCTAATGTTTATTTAAAAATGCAATTTTCCA	169636
repeat_region				Db	TTTTTAAAAAATGATTTCTTATGTCAGTTTCTAATGTTTATTTAAAAATGCAATTTTCCA	1320
repeat_region				Qy	TTTTTAAAAAATGATTTCTTATGTCAGTTTCTAATGTTTATTTAAAAATGCAATTTTCCA	169636
repeat_region				Db	TTTTTAAAAAATGATTTCTTATGTCAGTTTCTAATGTTTATTTAAAAATGCAATTTTCCA	1380
repeat_region				Qy	ATCTGTTACGCTCTGATGATGCTCTTTTAAATAAATAATATTTATTTGTTGGCTT	169576
repeat_region				Db	ATCTGTTACGCTCTGATGATGCTCTTTTAAATAAATAATATTTATTTGTTGGCTT	1440

Db	169575	ATCTGTTACGCGCTCTAGATGCGCTCTTTTAAATAATAATATTATTGTTGGCTT	169516
Qy	1441	TTGCTA-TTTTCTTTCTTATGCTCTTGAGAGGATATGTCAAATTTACTAGTGATCCA	1499
Db	169515	TTGCTATTTTTTTCTTATGCTCTTGAGAGGATATGTCAAATTTACTAGTGATCCA	169456
Qy	1500	AGAATAAACTTTGGCGTTGGCAATCTTTTCTCATCTATCTTTTCTGCTTATATTTATAA	1559
Db	169455	AGAATAAACTTTGGCGTTGGCAATCTTTTCTCATCTATCTTTTCTGCTTATATTTATAA	169396
Qy	1560	TTCTGTTCTGTTTATAATGCTCTTTTATCTTTCTTTGTTGTTACTTTGCTGTTCTTTT	1619
Db	169395	TTCTGTTCTGTTTATAATGCTCTTTTATCTTTCTTTGTTGTTACTTTGCTGTTCTTTT	169336
Qy	1620	GTAATACTCTCAGTAGAAGCTTAACCTTATGACATTCAGTCTTTCTCATTTCTACTAT	1679
Db	169335	GTAATACTCTCAGTAGAAGCTTAACCTTATGACATTCAGTCTTTCTCATTTCTACTAT	169276
Qy	1680	GAGTATTAGAGCCATAAAATTTCCCTTTAACTTCCCTTTCCACTTCAACTACATCTCAC	1739
Db	169275	GAGTATTAGAGCCATAAAATTTCCCTTTAACTTCCCTTTCCACTTCAACTACATCTCAC	169216
Qy	1740	AAATTTGGATTAGAGTAGTTTAAATTAATCATAGTATCTAAATTTTTTAAATTTCTGTA	1799
Db	169215	AAATTTGGATTAGAGTAGTTTAAATTAATCATAGTATCTAAATTTTTTAAATTTCTGTA	169156
Qy	1800	TTTTCTCTTTGATCCTGCAACTATTACAAAGTATTTTAAATCTGGAATAAAGAT	1859
Db	169155	TTTTCTCTTTGATCCTGCAACTATTACAAAGTATTTTAAATCTGGAATAAAGAT	169096
Qy	1860	TGTTATTTGTTATTTGTTGATCTGATCTCTAAATTTGAATATATTGAGATCAGATAATGTG	1919
Db	169095	TGTTATTTGTTATTTGTTGATCTGATCTCTAAATTTGAATATATTGAGATCAGATAATGTG	169036
Qy	1920	GTTTGTAGGACACTAATCCTTTGACAAATTTGTTGAGGCTTCTTTGGAACTTAATATGTC	1979
Db	169035	GTTTGTAGGACACTAATCCTTTGACAAATTTGTTGAGGCTTCTTTGGAACTTAATATGTC	168976
Qy	1980	TCAATTTTATAGACGTTCTGTGTTCTTTGGGAAAAACATGATTTTATGATGTTGTTGG	2039
Db	168975	TCAGTTTATAGACGTTCTGTGTTCTTTGGGAAAAACATGATTTTATGATGTTGTTGG	168916
Qy	2040	TTTAAATTTTGTATTTGTACATTTGATTTGCTTATTTATTTGGCTGAATCTCCA	2099
Db	168915	TTTAAATTTTGTATTTGTACATTTGATTTGCTTATTTATTTGGCTGAATCTCCA	168856
Qy	2100	TTATCCTTAATGCTCTCTCAATTTTGTCTGCTTCTTTTATTAATTAGAGATAATGTTA	2159
Db	168855	TTATCCTTAATGCTCTCTCAATTTTGTCTGCTTCTTTTATTAATTAGAGATAATGTTA	168796
Qy	2160	AATTTATCTCACCTCACATATGATGCTGTTTATATCTATATATAAATAATTTATAAT	2219
Db	168795	AATTTATCTCACCTCACATATGATGCTGTTTATATCTATATATAAATAATTTATAAT	168736
Qy	2220	CCATAAATTTATGTTATGTAATAATTTGGAGACCTTATATCATATATAAACAAGATGTTG	2279
Db	168735	CCATAAATTTATGTTATGTAATAATTTGGAGACCTTATATCATATATAAACAAGATGTTG	168676
Qy	2280	ATGAATGACAGACTTATCTATATGATAGTACCTTTTATCTCGTCATATGTTATTTG	2339
Db	168675	ATGAATGACAGACTTATCTATATGATAGTACCTTTTATCTCGTCATATGTTATTTG	168616
Qy	2340	ACTTTGCTCTAAATTTTTTTTAAATAATTTGTTGTTGTTATTTCTTTTTCAGCGGTTT	2399
Db	168615	ACTTTGCTCTAAATTTTTTTTAAATAATTTGTTGTTGTTATTTCTTTTTCAGCGGTTT	168556
Qy	2400	ATGTCAGTCTGTTGCAATGGTGACACAGTGAATTTATTTAGACATGTCAGCTTTTAA	2459
Db	168555	ATGTCAGTCTGTTGCAATGGTGACACAGTGAATTTATTTAGACATGTCAGCTTTTAA	168496
Qy	2460	TTATTTCTTTTCCAAATTTCTTTTAAATTTCTGATATACAAATTTTAGTCACTTTT	2519

Db

168495

TTATTTCTTTTCCATTTTCAATTTTATAAATTTCTGATATACAATATTATTAGTCACTTTT

168436

Qy

2520

ACCTTCTCTAGTGGAATTTT

2541

Db

168435

ACCTTCTCTAGTGGAATTTT

168414

RESULT 5

AP003400

135637 bp DNA linear PRI 15-MAR-2003

LOCUS

Homo sapiens genomic DNA, chromosome 11q, clone:RP11-643G5,

DEFINITION

complete sequence.

ACCESSION

AP003400

VERSION

AP003400.2

GI:15320508

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Published Only in Database (2001)

REFERENCE

2

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Direct Submission

COMMENT

Submitted (08-MAR-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

FEATURES

Location/Qualifiers

1..135637

organism="Homo sapiens"

mol_type="genomic DNA"

db_xref="taxon:9606"

chromosome="11q"

map="11q"

clone="RP11-643G5"

ORIGIN

Query Match 93.2%; Score 2367; DB 9; Length 135637;

Best Local Similarity 96.8%; Pred.No. 0;

Matches 2470; Conservative 0; Mismatches 70; Indels 11; Gaps 5;

Qy

1

AGCATAATTTTCTTTTAACTTACCTAGTGATTCAGATGAGTTCAATTTCTCATGTGAAT

60

Db

132900

AGCATAATTTTCTTTTAACTTACCTAGTGATTCAGATGAGTTCAATTTCTCATGTGAAT

132959

Qy

61

CACAGAAAAAATATGGGAAATTTGGAATATGCGGATGTGGCAAGTCGTAGTTGATTT

120

Db

132960

CACAGAAAAAATATGGGAAATTTGGAATATGCGGATGTGGCAAGTCGTAGTTGATTT

133019

Qy

121

GTTTCAATTTTAGCTTCTATCCATGTCAGAAAAGGAAATACTACTCTAGTCACGTA

180

Db

133020

GTTTCAATTTTAGCTTCTATCAGTGTGAGAAAAGGAAATACTACTCTAGTCACGTA

133079

Qy

181

CAATGTAATAATCTCCAGATCCCTCTCTCATCTCCAAATCCCTCCCAATTTGTAATCTG

240

Db

133080

CAATGTAATAATCTCCAGATCCCTCTCTCATCTCCAAATCCCTCCCAATTTGTAATCTG

133139

Qy

241

CACAGCAACACCTATACATATAATGTTGAAAGAAAGTTTGTATAGCAGATATCCCTTA

300

Db

133140

CTCTGGCCACACCTATACATATAATGTTGAAAGAAAGTTTGTATAGCAGATATCCCTTA

133199

Qy

301

ACTTTCTTCTGCTCTGAGGAAATACCTTTAAATTTCTTCTCACATTAATATTGGTATT

360

Db

133200

ACTTTCTTCTGCTCTGAGGAAATACCTTTAAATTTCTTCTCACATTAATATTGGTATT

133259

Qy

361

TACTGTTTTCTTTTCTGTTATCTTAAAGTACATCCCTTCCACATATACATAAATTTTTT

420

Db 133260 TACGTTTTCTTTTCGTTTATCTTAAGTACATCCCTTCCACATATACATAAATTTT 133319
QY 421 TGCTGTATGATTTATATAACACATTTTC-TTTAAGATGAGCATGCGATTTTCTTT 479
Db 133320 TGCTATATGATTTATATAACACATTTTCTTTGAGATGAGCAGGAGATTTTCTTT 133379
QY 480 CTTGATCTGAAGATGACATTCAGTGAATATCTAAATGTTAGAAATATCTTATTTTGA 539
Db 133380 CTTGATCTGAAGATGACATTCAGTGAATATCTAAATGTTAGAAATATCTTATTTTGA 133439
QY 540 ACTGTAAATATAGCATTAGGTTTAAATTCATATATGATATATATATGTTTATCT 599
Db 133440 ACTGTAAATATAGCATTAGGTTTAAATTCATATATGATATATATATGTTTATCT 133499
QY 600 CATAGCATGTTTACTGCTGTTTTCGAAATCTATTTGTTGGTTTAAATTTCAATATCT 659
Db 133500 CATAGCATGTTTACTGCTGTTTTCGAAATCTATTTGTTGGTTTAAATTTCAATATCT 133559
QY 660 ACTTGCTCTTAAATGAGTGGTGTTTTAAATATTTTCTGAAGTAGGTTTATGCAATT 719
Db 133560 ACTTGCTCTTAAATGAGTGGTGTTTTAAATATTTTCTGAAGTAGGTTTATGCAATT 133619
QY 720 AAATATTTTCTTAACTTAACTTCAACTCAAGGAAACAGTTGGCTTGACTCTGTT 779
Db 133620 AAATATTTTCTTAACTTAACTTCAACTCAAGGAAACAGTTGGCTTGACTCTGTT 133679
QY 780 TGTGGAAATTTTAAACTACTGTTTAAATTTCTTATTTGTTGTAATGACTATTTTAC 839
Db 133680 TGTGGAAATTTTAAACTACTGTTTAAATTTCTTATTTGTTGTAATGACTATTTTAT 133739
QY 840 GTCATATAACAAATTTTATTTGTTGTTAAATGACTTATTTGTTGTCATATGATAATTT 899
Db 133740 GTCATATAACAAATTTTATTTGTTGTTAAATGACTTATTTGTTGTCATATGATAATTT 133799
QY 900 ATGTCATAGAACAAATTTTATGCTGATATATGACTTATTTGTTATATGCTATACAAC 959
Db 133800 ATGTCATAGAACAAATTTTATGTTGATATATGACTTATTTGTTATATGCTATACAAC 133859
QY 960 TAGATTTTGTGTTGTTTACCGAGCTTTACTCTGTCACCCAGGCTGGAGTGTAATGG 1019
Db 133860 TAGATTTTGTGTTGTTTACCGAGCTTTACTCTGTCACCCAGGCTGGAGTGTAATGG 133919
QY 1020 CATGGCTCAGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCAATTTCTCCGCTCAGC 1079
Db 133920 CATGGCTCAGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCAATTTCTCCGCTCAGC 133979
QY 1080 CTCCAGTAGCTGGGACTACAGCATGAGCCACCGCAGCTAATTTTGTATTTT 1139
Db 133980 CTCCAGTAGCTGGGACTACAGCATGAGCCACCGCAGCTAATTTTGTATTTT 134039
QY 1140 AGTAGAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCTTGTAAATC 1199
Db 134040 AGTAGAGACGTGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCCTGACCTGTAATC 134099
QY 1200 CACCCGCTCGGCTGCAAGTGTGGGATTAAGCCAGTGGCCATGTCCTGGCCGA 1259
Db 134100 CACCCGCTCGGCTGCAAGTGTGGGATTAAGCCAGTGGCCATGTCCTGGCCGA 134159
QY 1260 TTTTAAATAAATGATTTCTTATGTCAGTTTTCATAGTTTATTTAAATGCAATTTCC 1319
Db 134160 TTTTAAATAAATGATTTCTTATGTCAGTTTTCATAGTTTATTTAAATGCAATTTCC 134219
QY 1320 ATTTGATGTAAGCTTTCAAAATTTATAGTATGTTGTTCTAGTATTTCTTATCTTTTGT 1379
Db 134220 ATTTGATGTAAGCTTTCAAAATTTATAGTATGTTGTTCTAGTATTTCTTATCTTTTGA 134279
QY 1380 AATCTGTACGCTGTAGATGCTCTTTTAAATAAATAATATTTATTTGTTGCGCT 1439
Db 134280 AATCTGTACGCTGTAGATGCTCTTTTAAATAAATAATATTTATTTGTTGCT 134339
QY 1440 TTTGCT--TATTTTCTTATGCTCTTGAGAGGATATGTCAAATTTACTAGTGATC 1497
Db 134340 TTTGCTATATTTTCTTATGCTCTTGAGAGGATATGTCGAATTTACTAGTGATC 134399
QY 1498 CAAAGATAAATTTTGGCTTTGGCAATCTTTTCTCACTATCTTTTGTCTTTATATTTTAT 1557
Db 134400 CAAAGATAAATTTTGGCTTTGGCAATCTTTTCTCACTATCTTTTGTCTTTATATTTTAT 134459
QY 1558 AATCTCTTCTTGTGTTTATATTAATTTGCTCTTTTATCTTTTGTGTTTCTTGTCTTCT 1617
Db 134460 AATCTCTTCTTGTGTTTATATTAATTTGCTCTTTTATCTTTTGTGTTTCTTGTCTTCT 134519
QY 1618 TTTGAAATCTCAGTGAAGTCTTAACTATTTGACATTCAGTCTTCTTCTTCTTCTTCTACT 1677
Db 134520 TTTGAAATCTCAGTGAAGTCTTAACTATTTGACATTCAGTCTTCTTCTTCTTCTTCTACT 134579
QY 1678 ATGATATTTAGAGCCATAAATTTCCCTTTAACTTCCCTTTCCACTTCAACTACATCTC 1737
Db 134580 ATGAGCAATTTAGAGCCATAAATTTCCCTTTAACTTCCCTTTCCACTTCAACTACATCTC 134639
QY 1738 ACAAATTTGAGATAGGAGTAGTTTAAATATCATTTAGTATCTAAATATTTTAAATTTCTG 1797
Db 134640 ACAAATTTGAGATAGGAGTAGTTTAAATATCATTTAGTATCTAAATATTTTAAATTTCTG 134699
QY 1798 TATTTCTTCTTGTGATCTGATCTCTAAATTTGAATATATTTGAGATCAGATATG 1857
Db 134700 TATTTCTTCTTGTGATCTGATCTCTAAATTTGAATATATTTGAGATCAGATATG 134758
QY 1858 ATGTTATTTGTTTGTGATCTGATCTCTAAATTTGAATATATTTGAGATCAGATATG 1917
Db 134759 ATGTTATTTGTTTGTGATCTGATCTCTAAATTTGAATATATTTGAGATCAGATATG 134818
QY 1918 TGGTTGTAGGACATAATCTTTTGACAAATTTGTTGAGGCTTCTTTTGGAACTTAAATATG 1977
Db 134819 TGGTTGTAGGACATAATCTTTTGACAAATTTGTTGAGGCTTCTTTTGGAACTTAAATATG 134878
QY 1978 GCTCAATTTTATAGAGCTTCTGTTTCTTTGGGAAAAACATGATTTGATGTTGTTT 2037
Db 134879 GCTCAATTTTATAAATGTTCTGTTTCTTTGGGAAAAACATGATTTGATGTTGTTT 134938
QY 2038 GGTTTAATATTTGTTATTTGATGATTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 2097
Db 134939 GGTTTAATATTTTATTTGATGATTTGATGATTTGATGATTTGATGATTTGATGATTT 134998
QY 2098 CATTTCTTAAATGTTGCTCTCTCATTTTGTCTCTCTTTTATTAATTTAGAGATAAATG 2157
Db 134999 CATTTCTTAAATGTTGCTCTCTCATTTTGTCTCTCTTTTATTAATTTAGAGATAAATG 135058
QY 2158 TAAATTTCTCCTCCTACTATAGTATGTTGTTTATATCTATATATATATAAATTTATAA 2217
Db 135059 TAAATTTCTCCTCCTACTATAGTATGTTGTTTATATCTATATATATATAAATTTATAA 135118
QY 2218 TTTCCATAAATTTTATGTTATGTTATTAATTTGGAGACCTTATATCATATATAAATCTAC 2271
Db 135119 TTTCCATAAATTTTATGTTATGTTATTAATTTGGAGACCTTATTAACATATATAAATAATAG 135178
QY 2272 AATTTGTATGAAATGACAGACTTATCTATGTTAGTAGGCTTTTATCTCTGCTATAAT 2331
Db 135179 AATAGTTGATGAAATGACAGACTTATCTATGTTAGTAGGCTTTTATCTCTGCTATAAT 135238
QY 2332 GTTATTTGATTTGCTTAAATTTTCTTAAATTTTCTTAAATTTTCTTAAATTTTCTTCTT 2390
Db 135239 GTTATTTGATTTGCTTAAATTTTCTTAAATTTTCTTAAATTTTCTTAAATTTTCTTCTT 135298
QY 2391 AGCGGTTTATGTCAGTCTGTTGCAATTTGTCACACAGCTGATTTTATTTAGACATCTAC 2450
Db 135299 AGTAGGTTTATGTCAGTCTGTTGCAATTTGTCACACAGCTGATTTTATTTAGACATCTAC 135358
QY 2451 GCTTTTAAATTTCTTTTCTTAAATTTTCTTAAATTTTCTTAAATTTTCTTAAATTTTCTT 2510
Db 135359 GCTTTTAAATTTCTTTTCTTAAATTTTCTTAAATTTTCTTAAATTTTCTTAAATTTTCTT 135418
QY 2511 GTCACTTTACCTTCTCTAGTGTGAATTT 2541
Db 135419 GTCACTTTACCTTCTCTAGTGTGAATTT 135449


```
RESULT 6
LOCUS AP003122 137888 bp DNA linear PRI 15-MAR-2003
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-31312,
complete sequence.
ACCESSION AP003122
VERSION AP003122.2 GI:15320502
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 137888)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Aug 27, 2001 this sequence version replaced gi:12597178.
FEATURES
source Location/Qualifiers
1..137888
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-31312"
ORIGIN
Query Match 93.2%; Score 2367; DB 9; Length 137888;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 70; Indels 11; Gaps 5;
QY 1 AGCATAAATTTTCTTTAACTTCTACCTAGTGAATTCAGATGAGTTCATTTCTCATGTGAAT 60
DB 104 AGCATAAATTTTCTTTAACTTCTACCTAGTGAATTCAGATGAGTTCATTTCTTAGTGAAT 163
QY 61 CACAGAAAATAATCGGAAATTTGGAATATGTGGGATGTGGCAAGTCGTAGTTGATTT 120
DB 164 CACAGAAAATAATCGGAAATTTGGAATATGTGGGATGTGGCAAGTCGTAGTTGATTT 223
QY 121 GGTTCATTTTAGCTTCTATCATGTGCAGAAAGGGAATAACTACTGTCTAGTCAGTA 180
DB 224 GGTTCATTTTAGCTTCTATCATGTGCAGAAAGGGAATAACTACTGTCTAGTCAGTA 283
QY 181 CAATGTAATAATGCGCCAGATCCCTTCATCTCAAAATCCCTCCCATTTGATTAATCG 240
DB 284 CAATGTAATAATGCGCCAGATCCCTTCATCTCAAAATCCCTCCCATTTGATTAATCG 343
QY 241 CACTGACCAAGACCTATACATATATGTTGAAAAGAGTTTGTAGCAGATATCCTTA 300
DB 344 CTTGCGCAAGACCTATACATATATGTTGAAAAGAGTTTGTAGCAGATATCCTTA 403
QY 301 ACTTTCTTCTGGTCTTTGAAGAAATACCTTTAATCTTTCTCAACCAATTAATTATGTAAT 360
DB 404 ACTTTCTTCTGGTCTTTGAAGAAATACCTTTAATCTTTCTCAACCAATTAATTATGTAAT 463
QY 361 TACTGTTTCTTTTCTGGTATCTTAAAGTACATCCCTTCCACATATACATAAAATTTT 420
DB 464 TACTGTTTCTTTTCTGGTATCTTAAAGTACATCCCTTCCACATATACATAAAATTTT 523
QY 421 TGCTGTATGATTTATATAAACAACATTTTTC--TTTAAAGATGAGTCGATTTTCTTT 479
```

524 TCGTATATGGATTATATATAAACAACATTTTCTTTTGAGATGAGCAGGAGATTTTCTTT 583

480 CTTGATCTGACAGATGACATTCAGTGTATTTCTAAATGTTAGAATAATCTTTATATTTTGA 539

584 CTTGATCTGACAGATGACATTCAGTGTATTTCTAAATGTTAGAATAATCTTTATATTTTGA 643

540 ACTTGTAATATTAGCATTTAGGCTTTTAAATTCATATATTGATATATATTATGTTTATCT 599

644 ACTTGTAATATTAGCATTTAGGCTTTTAAATTCATATATTGATATATATTATGTTTATCT 703

600 CATAGCATGTTTAGTCTGTGTTTTCATCTATTGTTGGGTATAAATTCATATTTCT 659

704 CATAGCATGTTTAGTCTGTGTTTTCATCTATTGTTGGGTATAAATTCATATTTCT 763

660 ACTTGCTTTCTAAATAGAGTTCGGTTTTAAATTTCTTGAAGTAGGTTTATTTGCAATT 719

764 ACTTGCTTTCTAAATAGAGTTCGGTTTTAAATTTCTTGAAGTAGGTTTATTTGTTAAAT 823

720 AAATATTATTTTCTTTAACTTTTCAAACTCAAGGAAACCCAGTTGGCTTGCCTCTGTT 779

824 AAATATTATTTTCTTTAACTTTTCAAACTCAAGGAAACCCAGTTGGCTTGCCTCTGTT 883

780 TGTGGAATAATTTTAAACTACTGTTTAAATTTCTTTATGTTGTTGTAATATGACTATTTTAC 839

884 TGTGGAATAATTTTAAACTACTGTTTAAATTTCTTTATGTTGTTGTAATATGACTATTTTAT 943

840 GTCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAAATTTT 899

944 GTCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAAATTTT 1003

900 ATGTCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTATATGGCTATATCAAC 959

1004 ATGTCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTATATGGCTATATCAAC 1063

960 TAGATTTTCTGTTGTTTGTACCGAGTCTTACTCTGTCAACCCAGGCTCGAGTGTAAATGG 1019

1064 TAGATTTTCTGTTGTTTGTACCGAGTCTTACTCTGTCAACCCAGGCTCGAGTGTAAATGG 1123

1020 CATGCTCTCAGCTCACTGCAACCTCGCTCCCGGTTCAAGCCATTTTCCACCTCAGC 1079

1124 CATGCTCTCAGCTCACTGCAACCTCGCTCCCGGTTCAAGCCATTTTCCCGCTCAGC 1183

1080 CTCCCAAGTAGCTGGACCTACAGGCATGAGCCACCGCACCGGCTAAATTTTGTATTTT 1139

1184 CTCCCAAGTAGCTGGACCTACAGGCATGAGCCACCGCACCGGCTAAATTTTGTATTTT 1243

1140 AGTAGAGAGCTGTTCCACTATGTTGGCCAGGCTGATCTCGAACTCTCGACCTTGTAAATC 1199

1244 AGTAGAGAGCTGTTCCACTATGTTGGCCAGGCTGCTCTCGAACTCTGAACTCGTAATC 1303

1200 CACCGGCTCGGCTCGCCAAAGTGTGGATTAACAGGCTGAGCCATGTCCTGCGCCGA 1259

1304 CACCGGCTCGGCTCGCCAAAGTGTGGATTAACAGGCTGAGCCATGTCCTGCGCCGA 1363

1260 TTTTAAATAATGATTTCTTATGTCAGTTTTCATAGTTTATTTTAAATGCAATTTCC 1319

1364 TTTTAAATAATGATTTCTTATGTCAGTTTTCATAGTTTATTTTAAATGCAATTTCC 1423

1320 ATTTGATGTAAGCTTTCAAAATTTATAGTATAGTTTCTCTAGTATTTTCTTATCTTTTGT 1379

1424 ATTTGATGTAAGCTTTCAAAATTTATAGTATAGTTTCTCTAGTATTTTCTTATCTTTTAT 1483

1380 AATCTGTTCAAGCTCTGTAGATGTGCTCTTTTAAATAAATAATATATTTGTTGCTGCT 1439

1484 AATCTGTTCAAGCTCTGTAGATGTGCTCTTTTAAATAAATAATATATTTGTTGCTGCT 1543

1440 TTTGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1497

1544 TTTGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1503

1498 CAAAGAAATAAATCTTTGGCTTTGGCAATCTTTCTCATCTATCTTTGCTTTTATTTTAT 1557

1604 CAAAGAAATAAATCTTTGGCTTTGGCAATCTTTCTCATCTATCTTTGCTTTTATTTTAT 1663

QY 720 AATTATATTTTCTTCTTAACTTTCAAACTCAAGGAAACAGTTGGCTTGACTCTGTT 779
DB 174245 AATTATATTTTCTTCTTAACTTTCAAACTCAAGGAAACAGTTGGCTTGACTCTGTT 174186
QY 780 TGTGGAATAATTTAACTACTGTTTAAATTTCTTTATTTGGTTGTAATAGACTATTTTAC 839
DB 174185 TGTGGAATAATTTAACTACTGTTTAAATTTCTTTATTTGGTTGTAATAGACTATTTTAT 174126
QY 840 GTCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTT 899
DB 174125 GTCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTT 174066
QY 900 ATGTCATAGACAAATTTTATTTGTTGTCATATGACTTTATTTGTTGTCATATGACTATTAAC 959
DB 174065 ATGTCATAGACAAATTTTATTTGTTGTCATATGACTTTATTTGTTGTCATATGACTATTAAC 174006
QY 960 TAGATTTTTTTTGTGTTTGTGACCGAGCTTACTCTGTCACCCAGGCTGGAAGTGAATGG 1019
DB 174005 TAGATTTTTTTGTTTGTGACCAAGCTTACTCTGTCACCCAGGCTGGAAGTGAATGG 173946
QY 1020 CATGGTCTCAGCTCACTGCAACCTCCGCTCCCGGTTTCAAGCCATTTCTCCACCTCAGC 1079
DB 173945 CATGGTCTCAGCTCACTGCAACCTCCGCTCCCGGTTTCAAGCCATTTCTCCCGCTCAGC 173886
QY 1080 CTCCTCAAGTACCTGGGACTACAGGATGAGCCACCGCACCGGCTAATTTTGTATTTT 1139
DB 173885 CTCCTCAAGTACCTGGGACTACAGGATGAGCCACCGCACCGGCTAATTTTGTATTTT 173826
QY 1140 AGTAGAGACCTGTTTCCACTATGTTGGCCAGGCTGATCTCGAACTCTCGACTGTTGTAATC 1199
DB 173825 AGTAGAGACCTGTTTCCACTATGTTGGCCAGGCTGATCTCGAACTCTCGACTGTTGTAATC 173766
QY 1200 CACCGGCTCGGCTGCTCAAGTCTGGGATACAGCGTGAAGCCATTTGTCCTGGCCGA 1259
DB 173765 CACCGGCTCGGCTGCTCAAGTCTGGGATACAGCGTGAAGCCATTTGTCCTGGCCGA 173706
QY 1260 TTTTAAATAAATGATTTCTATGCTAGTTTCTATAGTTTATTTAAATGCAATTTTC 1319
DB 173705 TTTTAAATAAATGATTTCTATGCTAGTTTCTATAGTTTATTTAAATGCAATTTTC 173646
QY 1320 AATTGATGTAAGCTTTCAAAATTTATAGTATAGTTGTTCTAGTATTTCTTATCTTTGTT 1379
DB 173645 AATTGATGTAAGCTTTCAAAATTTATAGTATAGTTGTTCTAGTATTTCTTATCTTTGTT 173586
QY 1380 AATCTGTTGAGCTCTGATAGTGGCTCTTTTAAATAAATAATTTATTTGTTGCGCT 1439
DB 173585 AATCTGTTGAGCTCTGATAGTGGCTCTTTTAAATAAATAATTTATTTGTTGCGCT 173526
QY 1440 TTTGCT--TATTTTCTTATTTGCTCTGAGGAGGATATGCAATTTACTAGTATATC 1497
DB 173525 TTTGCTATATTTTCTTATTTGCTCTGAGGAGGATATGCAATTTACTAGTATATC 173466
QY 1498 CAAAGAATAAATTTGGCGTTGGCAATCTTTCTCATCTATCTTTGCTTATATTTTAT 1557
DB 173465 CAAAGAATAAATTTGGCGTTGGCAATCTTTCTCATCTATCTTTGCTTATATTTTAT 173406
QY 1558 AATCTGTTGCTGTTTAAATGCTCTTTTATCTTTTGTGTTTACTTTGCTGTTCT 1617
DB 173405 AATCTGTTGCTGTTTAAATGCTCTTTTATCTTTTGTGTTTACTTTGCTGTTCT 173346
QY 1618 TTGTAAATCTCAGTAGAATGCTTAACTATTTGACATTCAGCTTTCTTCTTCTACT 1677
DB 173345 TTGTAAATCTCAGTAGAATGCTTAACTATTTGACATTCAGCTTTCTTCTTCTACT 173286
QY 1678 ATGAGTATTTAGAGCCATAATTTCCCTTTAACTTTCCCTTCCCTTCCCTCACTGACTATCTC 1737
DB 173285 ATGAGTATTTAGAGCCATAATTTCCCTTTAACTTTCCCTTCCCTTCCCTCACTGACTATCTC 173226
QY 1738 ACAATTTGAGTAGGATGTTTAAATFATCATAGTATCTAAATATTTTAAATTTCTG 1797
DB 173225 ACAATTTGAGTAGGATGTTTAAATFATCATAGTATCTAAATATTTTAAATTTCTG 173166
QY 1798 TATTTCTCTTGTGCTGCAACTATTTTCAAGTATTTTAAATCTGCAATATAAAG 1857

DB 173165 TATTTCTCTTCTGATCTGGAACCTGTTTACAAG-ATTTTTTAAATCTGTAATGTAAG 173107
QY 1858 AATCTTATGTTATTTGTTGATCTGATCTCTAAATGCAATATTTGAGTACAGATAATG 1917
DB 173106 AATCTTATGTTATTTGTTGATCTGATCTCTAAATGCAATATTTGAGTACAGATAATG 173047
QY 1918 TGGTTTGTAGGACACTAATCTTTGACAATTTGAGGCTTCCCTTTGGAACCTAATATG 1977
DB 173046 TGGTTTGTAGGACACTAATCTTTGACAATTTGAGGCTTCCCTTTGGAACCTAATATG 172987
QY 1978 GCTCAATTTTATAGAGCTTCTGTTTCTTTGGGAAAAACATGATTTATGATGTTGTT 2037
DB 172986 GCTCAATTTTATAAATGTTCTGTTTCTTTGGGAAAAACATGATTTATGATGTTGTT 172927
QY 2038 GGTTTAATATTTTCTTATTTGACATTTAGTTTGGCTTTGCTTATTTATTTGGCTGAAATCTC 2097
DB 172926 GGTTTAATATTTTATTTTGTACATTTAGTTTGGCTTTGCTTATTTATTTGGCTGAAATCTC 172867
QY 2098 CATTTATCTTAAATGCTCTCTCATTTTGTCTGCTTCTTTTATTTAAATTTAGAGATAAATG 2157
DB 172866 CATTTATCTTAAATGCTCTCTCATTTTGTCTGCTTCTTTTATTTAAATTTAGAGATAAATG 172807
QY 2158 TAAATTTATCTCACTCACTATAGTGTCTGTTTATCTATATATATATAAATTTATAA 2217
DB 172806 TAAATTTATCTCACTCACTATAGTGTCTGTTTATCTATATATATATAAATTTATAA 172747
QY 2218 TTTCCATAAATTTATGTTATGTTATGTTTGGAGACTTATTAATATATAAATTTATAA 2271
DB 172746 TTTCCATAAATTTATGTTATGTTTATGTTTGGAGACTTATTAATATATAAATTTATAA 172687
QY 2272 AATTTGTTGATGAAATGACAGACTTATATCTTATGTTAGTGGCTTTTATCTGTCATAAT 2331
DB 172686 AATTTGTTGATGAAATGACAGACTTATATCTTATGTTAGTGGCTTTTATCTGTCATAAT 172627
QY 2332 GTTATTTGACTTGTCTCTAAATTTTTTTTTTAAATTTATTTGTTGTTGTTTCTTTTTC 2390
DB 172626 GTTATTTGACTTGTCTCTAAATTTTTTTTTTAAATTTATTTGTTGTTGTTTCTTTTTC 172567
QY 2391 AGCGGTTTATGTCAGTCTGTTCAATTTGTTAGTGGCTTATTTATTTAGACATGCTAC 2450
DB 172566 AGTGGTTTATGTCAGTCTGTTCAATTTGTTAGTGGCTTATTTATTTAGACATGCTAC 172507
QY 2451 GCTTTTAAATTTCTTTTCTTATTTTCAATTTTCTTATATTTCTGATATCAATATTTAG 2510
DB 172506 GCTTTTAAATTTCTTTTCTTATTTTCAATTTTCTTATATTTCTAATATCAATATTTAG 172447
QY 2511 GTCATTTTACCTTCTCTAGTGTGAATTT 2541
DB 172446 GTCATTTTACCTTCTCTAGTGTGAATTT 172416

RESULT 8

AC023784

LOCUS

DEFINITION

AC023784

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AC023784 117521 bp DNA linear HTG 07-JUL-2000
Homo sapiens chromosome 11 clone CTD-2026G22, WORKING DRAFT
SEQUENCE, 23 unordered pieces.

AC023784 GI:7230809
HTG; HTGS_PHASE1; HTGS_DRAFT.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 117521)

Waterston,R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 117521)

Waterston,R.H.

Direct Submission

Submitted (17-FEB-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA
On Mar 13, 2000 this sequence version replaced gi:7021753.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.MS2026G22
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 108841 bases at least Q40
Consensus quality: 111448 bases at least Q30
Insert size: 132000; agarose-fp
Insert size: 115321; sum-of-contigs
Quality coverage: 3.18 in Q20 bases; agarose-fp
Quality coverage: 3.67 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1436: contig of 1436 bp in length
* 1437 1536: gap of unknown length
* 1537 2618: contig of 1082 bp in length
* 2619 2718: gap of unknown length
* 2719 4121: contig of 1403 bp in length
* 4122 4221: gap of unknown length
* 4222 5276: contig of 1055 bp in length
* 5277 5376: gap of unknown length
* 5377 6917: contig of 1541 bp in length
* 6918 7017: gap of unknown length
* 7018 8794: contig of 1777 bp in length
* 8795 8894: gap of unknown length
* 8895 13539: contig of 4645 bp in length
* 13540 13639: gap of unknown length
* 13640 17094: contig of 3455 bp in length
* 17095 17194: gap of unknown length
* 17195 20927: contig of 3733 bp in length
* 20928 21027: gap of unknown length
* 21028 24808: contig of 3781 bp in length
* 24809 24908: gap of unknown length
* 24909 28904: contig of 3996 bp in length
* 28905 29004: gap of unknown length
* 29005 32266: contig of 3262 bp in length
* 32267 32366: gap of unknown length
* 32367 36987: contig of 4621 bp in length
* 36988 37087: gap of unknown length
* 37088 40336: contig of 3249 bp in length
* 40337 40436: gap of unknown length
* 40437 46367: contig of 5931 bp in length
* 46368 52979: contig of 6512 bp in length
* 52980 53079: gap of unknown length
* 53080 58971: contig of 5892 bp in length
* 58972 59072: gap of unknown length
* 59073 65824: contig of 6753 bp in length
* 65825 65924: gap of unknown length
* 65925 72767: contig of 6843 bp in length
* 72768 72867: gap of unknown length
* 72868 80928: contig of 8061 bp in length
* 80929 81028: gap of unknown length
* 81029 92294: contig of 11266 bp in length
* 92295 92394: gap of unknown length

* 92395 103346: contig of 10952 bp in length
* 103347 103446: gap of unknown length
* 103447 117521: contig of 14075 bp in length.
FEATURES
Location/Qualifiers
1..117521
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="CTD-2026G22"

ORIGIN

Query Match 84.5%; Score 2148; DB 2; Length 117521;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 2179; Conservative 0; Mismatches 115; Indels 2; Gaps 2;
QY 247 CCAAGACCTATACAATAAATCTTGAAGAAGAGTTTTCATAGCAGATATCTTAACTTTC 306
Db |||||
QY 24788 CTAAGTCGTACTAAATCAAGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 24847
Db |||||
QY 307 TTCTGGTCTTGAAGGAATACCTTTAAATCTTCTCACCAATTAATATGATGTAATCTGT 366
Db NNN 24907
QY 367 TTCTCTTTCTCGTTATCTTAAAGTACATCCCTTCCACATATCTTAAATTTTTCCTGT 426
Db |||||
QY 24908 NTCTCTTTCTCGTTATCTTAAAGTACATCCCTTCCACATATCTTAAATTTTTCCTGT 24967
Db |||||
QY 427 ATGGATTATATTAACAACATTTTCTTTAAGATGAGCATGCGATTTTCTTTCTTGTATC 486
Db |||||
QY 24968 ATGGAATATATTAACAACATTTTCTTTAAGATGAGCATGCGATTTTCTTTCTTGTATC 25027
Db |||||
QY 487 TGAAGATGATCTAGTGATATCTTAATGTTTGAATAATCTTATATTTTGAACCTGTA 546
Db |||||
QY 25028 TGAAGATGATCTAGTG- TATCTAAATCTTGAATAATCTTATATTTTGAACCTGTA 25086
Db |||||
QY 547 ATATTAGCATTTAGGTTTAAATTCATATATTTGATATATTTATGTTTATCTCATAGCA 606
Db |||||
QY 25087 ATATTAGCATTTAGGTTTAAATTCATATATTTGATATATTTATGTTTATCTCATAGCA 25146
Db |||||
QY 607 TGTTAGTCTGTGTTTTCGAATCTATTTGTTGGGTTATAATTTCAATTTCTTCTTGTCC 666
Db |||||
QY 25147 TGTTAGTCTGTGTTTTCGAATCTATTTGTTGGGTTATAATTTCAATTTCTTCTTGTCC 25206
Db |||||
QY 667 TTCTAAATAGTTGGGTTTAAATTTCTGAAGTAGTTTATTTGAAATTTAAATTTAT 726
Db |||||
QY 25207 TTCTAAATAGTTGGGTTTAAATTTCTGAAGTAGTTTATTTGCAATTTAAATTTAT 25266
Db |||||
QY 727 TTTTCTTTTAACTTTCAAACTCAAGGAACCACTTGCCCTTGACTCTGTTTGTGGAA 786
Db |||||
QY 25267 TTTTCTTTTAACTTTCAAACTCAAGGAACCACTTGCCCTTGACTCTGTTTGTGGAA 25326
Db |||||
QY 787 AATTTTAACTACTGTTTAAATTTCTTATTTGTTGTTAATATGACTATTTTACGTCAAT 846
Db |||||
QY 25327 AATTTTAACTACTGTTTAAATTTCTTATTTGTTGTTAATATGACTATTTTACGTCAAT 25386
Db |||||
QY 847 AACAAATTTATTTGTTTAAATGACTTTTATTTGTTTGCATATGATATTTTATGTCAT 906
Db |||||
QY 25387 AACAAATTTATTTGTTTAAATGACTTTTATTTGTTTGCATATGATATTTTATGTCAT 25446
Db |||||
QY 907 AGAACAATTTTATTTGTTTAAATGACTTTTATTTGTTTGCATATGATATTTTACGTCAAT 966
Db |||||
QY 25447 AGAACAATTTTATTTGTTTAAATGACTTTTATTTGTTTGCATATGATATTTTACGTCAAT 25506
Db |||||
QY 967 TTTTGTGTTTGTGACCGAGTCTTACTCTGTCAACCGAGGTGAGGTGAATGCAATGTC 1026
Db |||||
QY 25507 TTTTGTGTTTGTGACCGAGTCTTACTCTGTCAACCGAGGTGAGGTGAATGCAATGTC 25566
Db |||||
QY 1027 TCAGCTCACTGCAACCTCCGCTCCGCGTTTGAAGCAATTTCTCCACCTCAGCCTCCCAA 1086
Db |||||
QY 25567 TCAGCTCACTGCAACCTCCGCTCCGCGTTTGAAGCAATTTCTCCACCTCAGCCTCCCAA 25626
Db |||||
QY 1087 GTAGCTGGGACTACAGGCATGAGCCACCGCAGCCGGCTAATTTTGTATTTTAGTAGAG 1146
Db |||||

```
Db 25627 GTAGCTGGACCTACAGGCAATGAGCCACCGCCCGCTAAATTTTGTATATTTTAGTAGAG 25686
Qy 1147 ACGTGGTTTCCATATGTTGGCCAGGCTGATCTCGAACTCTCTGACCTTGTATATCCACCGC 1206
Db 25687 ACGTGGTTTCCATATGTTGGCCAGGCTGATCTCGAACTCTCTGACCTTGTATATCCACCGC 25746
Qy 1207 CTCGGCTGCCAAAGTCTGGGATACAGGCGTGAGCCATGTGCGTGGCCGATTTTTTA 1266
Db 25747 CTCGGCTGCCAAAGTCTGGGATACAGGCGTGAGCCATGTGCGTGGCCGATTTTTTA 25806
Qy 1267 AAAAATGTAATCTTATGTCAGTTTTTCATAAGCTTTTATTAAGTCAATTTTCCATTTGAT 1326
Db 25807 AAAAATGTAATCTTATGTCAGTTTTTCATAAGCTTTTATTAAGTCAATTTTCCATTTGAT 25866
Qy 1327 GTAAGCTTTCAAAATTTATAGTATAGTTGTTCTCTAGTATTTTCTTATCTTTTGTATCTGT 1386
Db 25867 GTAAGCTTTCAAAATTTATAGTATAGTTGTTCTCTAGTATTTTCTTATCTTTTGTATCTGT 25926
Qy 1387 TCAGCGTCTGAGATGTCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1446
Db 25927 TCAGCGTCTGAGATGTCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 25986
Qy 1447 TTTTTCCTTATGCTTTGAGAGGATATGTCAAATTTTCTCTATCTATCTTTGCTTTATATTTTAAATCTGT 1565
Db 25987 TTTTTCCTTATGCTTTGAGAGGATATGTCAAATTTTCTCTATCTATCTTTGCTTTATATTTTAAATCTGT 26046
Qy 1506 AAATCTTGGCGTTGGCAATCTTTTCTCAATCTATCTTTGCTTTATATTTTAAATCTGT 1565
Db 26047 AAATCTTGGCGTTGGCAATCTTTTCTCAATCTATCTTTGCTTTATATTTTAAATCTGT 26106
Qy 1566 TCTTGTATTAATGCTCTTTTATCTTCTTGTGTTTACTTGTGTTTCTTGTGTTTCTTGTGTTTCT 1625
Db 26107 TCTTGTATTAATGCTCTTTTATCTTCTTGTGTTTACTTGTGTTTCTTGTGTTTCTTGTGTTTCT 26166
Qy 1626 TCTCTAGTAAAGCTTAACTTATGACATTCAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1685
Db 26167 TCTCTAGTAAAGCTTAACTTATGACATTCAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 26226
Qy 1686 TTAGAGCCATAAATTTCCCTTTAACTTCCCTTTCACTTCCACTCACTCACTCACTCACTCACT 1745
Db 26227 TTAGAGCCATAAATTTCCCTTTAACTTCCCTTTCACTTCCACTCACTCACTCACTCACTCACT 26286
Qy 1746 GGAATTAGAGTAGTAAATTAATCATGATGATCAATCAATCAATCAATCAATCAATCAATCAAT 1805
Db 26287 GGAATTAGAGTAGTAAATTAATCATGATGATCAATCAATCAATCAATCAATCAATCAATCAAT 26346
Qy 1806 TCTTTCATCTGCAACTTATTTACAAGTATTTTAAATCTTCAATCAATCAATCAATCAATCAAT 1865
Db 26347 TCTTTCATCTGCAACTTATTTACAAGTATTTTAAATCTTCAATCAATCAATCAATCAATCAAT 26406
Qy 1866 TGTATTTGTTGATCTGATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1925
Db 26407 TGTATTTGTTGATCTGATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 26466
Qy 1926 AGGACACTAATCTTTGCAATGTTGAGCTTCTTGGAACTTAAATTAATTAATTAATTAATTAAT 1985
Db 26467 AGGACACTAATCTTTGCAATGTTGAGCTTCTTGGAACTTAAATTAATTAATTAATTAATTAAT 26526
Qy 1986 TTTATAGAGCTTCTGTTTCTTTGGGAAACATGATTTGATGATGATGATGATGATGATGATGATGAT 2045
Db 26527 TTTATAGAGCTTCTGTTTCTTTGGGAAACATGATTTGATGATGATGATGATGATGATGATGATGAT 26586
Qy 2046 ATTTCGATTTGACATTAATGATTTGATTTGCTTTATTTGCTTTGCTTTGCTTTGCTTTGCTTTGCT 2105
Db 26587 ATTTCGATTTGACATTAATGATTTGATTTGCTTTATTTGCTTTGCTTTGCTTTGCTTTGCTTTGCT 26646
Qy 2106 TTAATGCTCTCTCATTTTGTCTCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2165
Db 26647 TTAATGCTCTCTCATTTTGTCTCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 26706
Qy 2166 CTCACCTCACTAGTATGCTGTTTATATATATATATATATATATATATATATATATATATATATATAT 2225
```

```
Db 26707 CTCACCTCACTATAGTATGCTGTTTATATATATATATATATATATATATATATATATATAT 26766
Qy 2226 ATTATATGTTATGATAATTTGGAGACCTATTTATCATATATATATATATATATATATATATATAT 2285
Db 26767 ATTATATGTTATGATAATTTGGAGACCTATTTATCATATATATATATATATATATATATATAT 26826
Qy 2286 TGACAGACTTATATCTTATGATAGTACCTTTTATCTCTGTCATATATGTTATTTGACTTTTG 2345
Db 26827 TGACAGACTTATATCTTATGATAGTACCTTTTATCTCTGTCATATATGTTATTTGACTTTTG 26886
Qy 2346 TCTTAAATTTTATTTTAAATTAATTTGTTGGTATTTCTTTTTCAGCGGTTTTATGTCAT 2405
Db 26887 TCTTAAATTTTATTTTAAATTAATTTGTTGGTATTTCTTTTTCAGCGGTTTTATGTCAT 26946
Qy 2406 CTGCTGTCGAATTTGTCACAGCTGATTTATTTAGCATGCTACGCTTTTATTTATTTATTTT 2465
Db 26947 CTGCTGTCGAATTTGTCACAGCTGATTTATTTAGCATGCTACGCTTTTATTTATTTATTTT 27006
Qy 2466 TTTTTCCTATTTTCAATTTTATATATTTCTGATATACAAATTTTAGTTCACCTTTTACCTTC 2525
Db 27007 TTTTTCCTATTTTCAATTTTATATATTTCTGATATACAAATTTTAGTTCACCTTTTACCTTC 27066
Qy 2526 CTCTAGTGTGAATTTT 2541
Db 27067 CTCTAGTGTGAATTTT 27082
```

RESULT 9
AC074003/c

LOCUS AC074003 246865 bp DNA linear HTG 24-SEP-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-163G14, *** SEQUENCING IN
PROGRESS ***, 53 unordered pieces.
ACCESSION AC074003 GI:10281455
VERSION HTG; HTGS_PHASE1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 246865)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 246865)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 24, 2000 this sequence version replaced gi:9186850.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

----- Project Information -----
Center project name: H_MH0163G14

----- Summary Statistics -----
Sequencing vector: M13; 100%

Sequencing vector: plasmid; 0%

Chemistry: Dye-terminator Big Dye; 0% of reads

Chemistry: Dye-terminator Big Dye; 0% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 213368 bases at least Q40

Consensus quality: 223527 bases at least Q30

Consensus quality: 229418 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 241665; sum-of-contigs

Quality coverage: 4.83 in Q20 bases; agarose-fp

Quality coverage: 3.37 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently

consists of 53 contigs. The true order of the pieces

misc_feature	111982..111324	/note="assembly_name:Contig59"	
misc_feature	113625..114983	/note="assembly_name:Contig62"	
misc_feature	115084..116145	/note="assembly_name:Contig63"	
misc_feature	116246..117625	/note="assembly_name:Contig64"	
misc_feature	117726..120009	/note="assembly_name:Contig65"	
Query Match	58.1%; Score 1476.6; DB 2; Length 246865;		
Best Local Similarity	96.7%; Pred No. 5.2e-218;		
Matches 1529; Conservative	0; Mismatches 49; Indels 3; Gaps 2;		
QY	964	TTTTTTTGTGTTTGTGACCGAGTCTTACTCTGTGACCCAGCGCTGGAGTGAATGGCATG	1023
DB	48628	TCTTTTTTTTTTTTGGAGATGGAGTCTCACTCTGTGCGCCAGCGTGGAGTGGCGG	48569
QY	1024	GTCTCAGCTCACTGCAACTCGGCTCGGGTTCAAGCCATCTTCCACCTCAGCCCTCC	1083
DB	48568	ATCTCGGCTCACTGCAACTCTGCGCTCGGGTTTCATACCACTCTCCGGCCTCAGCCCTCC	48509
QY	1084	CAAGTAGCTGGGACTACAGGATGAGCCACCGCACCGCGCTAA--TTTTTGTATTTTAG	1141
DB	48508	CGAGTAAGTGGGACTACAGGACCGCGCACCGCATGCTCGCTAATTTTTTGTATTTTAG	48449
QY	1142	TAGAGAGTGTTCACATPATGTTGGCCAGGTGATCTCGAACTCTCGACCTTGTAAATCCA	1201
DB	48448	TAGAGACAGGTTTCAACATGTTAAACAGGTGATCTCGAACTCTCGACCTCGTATCCA	48389
QY	1202	CCGCCCTCGGCTGCCAAAGTGTGGGATTAACAGCGTGAGCCATTTGCGTCGCGGAT	1261
DB	48388	CTCATGTGAGCTCCCAAGTGTGGGATTAACAGCGTGAGCCATTTGCGTCGCGGAT	48329
QY	1262	TTTTTAAAGATGATCTTATGTCAGTCTTCTAAGTTTCTAAGTTTATTTAAAGATTTCCAT	1321
DB	48328	TTTTTAAAGATGATCTTATGTCAGTCTTCTAAGTTTATTTAAAGATTTCCAT	48269
QY	1322	TTGATGTAAGCTTTCAAATTTATAGTATAGTGTGTTCTCTAGTATTTCTTATCTTTTGTAA	1381
DB	48268	GTGATGTAAGCTTTCAAATTTATAGTATAGTGTGTTCTCTAGTATTTCTTATCTTTTGTAA	48209
QY	1382	TCGTGTCAGGCTGTAGATGCGCTCTTTTAAATAAATAATATTTGTTGCGCTTT	1441
DB	48208	TCGTGTCAGGCTGTAGATGCGCTCTTTTAAATAAATAATATTTGTTGCGCTTT	48149
QY	1442	TGCTA--TTTTTTTTTCTTATGCTCTGAGAGGATATGTCAAATTTACTAGTGTATCCAA	1500
DB	48148	TGCTA--TTTTTTTTTCTTATGCTCTGAGAGGATATGTCAAATTTACTAGTGTATCCAA	48089
QY	1501	AGAAATAAATTTGGCGTGGCAATCTTTTCTCATCTATCTTTGCTTTATATTTTATTAAT	1560
DB	48088	AGAAATAAATTTGGCTTTGGCAATCTTTTCTCATCTATCTTTGCTTTATATTTTATTAAT	48029
QY	1561	TCGTGTCGTTTATTAATTTGCTCTTTTATCTTTGTTTACCTTTGTTTACCTTTGTTTTC	1620
DB	48028	TCGTGTCGTTTATTAATTTGCTCTTTTATCTTTTATCTTTGTTTACCTTTGTTTTC	47969
QY	1621	TAAATCTCTCAGTAGAATGCTTAACTATTATGACATTCAGCTTTCTCATTTCTACTATG	1680
DB	47968	TAAATCTCTCAGTAGAATGCTTAACTATTATGACATTCAGCTTTCTCATTTCTACTATG	47909
QY	1681	AGTATTAGAGCCATAAATTTCCCTTTTAACTTTCCCTTTCCCTTTCCCTTTCACTCATCACA	1740
DB	47908	AGTATTAGAGCCATAAATTTCCCTTTTAACTTTCCCTTTCCCTTTCCCTTTCACTCATCACA	47849
QY	1741	AATTTGATTTAGAGTAGTTTAAATATATATATATATATATATATATATATATATATATAT	1800
DB	47848	AATTTGATTTAGAGTAGTTTAAATATATATATATATATATATATATATATATATATATAT	47789
QY	1801	TTTTCTTTTGTGCTCGCACTTTTACAGTATTTTAAATTTTAAATTTTAAATTTTAAAGATT	1860
DB		TTTTCTTTTGTGCTCGCACTTTTACAGTATTTTAAATTTTAAATTTTAAATTTTAAAGATT	
Db	47788	TTTCTTCTTTGATCTGCAACTATTATCAAGTATTTTTTAAAAATCTGATATATAAGATT	47729
QY	1861	GTATTGTTTATTTGATCTGATCTCTAAATTCGAATATATTGATCAGATATATGCG	1920
DB	47728	GTATTGTTTATTTGTTGATCTGATCTCTAAATTCGAATATATTGATCAGATATATGCG	47669
QY	1921	TTTGTAGGACACTAATCTCTTTTGACAAATTTGAGGCTTCTTTTGGAACTAATATGCT	1980
DB	47668	TTTGTAGGACACTAATCTCTTTTGACAAATTTGAGGCTTCTTTTGGAACTAATATGCT	47609
QY	1981	CAATTTTATAGACGTTCTGCTGTTTCTTTGGGAAAAACATGATTTATGATGTTGTTGCT	2040
DB	47608	CAGTTTTTATAGACGTTCTGCTGTTTCTTTGGGAAAAACATGATTTATGATGTTGTTGCT	47549
QY	2041	TTAATATTGTTGTTGATTTGACATTTAGTTTGGTTGCTTTATTTTGGCTGAATCTCCAT	2100
DB	47548	TTAATATTGTTGTTGATTTGACATTTAGTTTGGTTGCTTTATTTTGGCTGAATCTCCAT	47489
QY	2101	TATCCTTAATGCTCTCTCAATTTTGTCTTCTTTTATTTAAATPAGAGATATAATGTTAA	2160
DB	47488	TATCCTTAATGCTCTCTCAATTTTGTCTTCTTTTATTTAAATPAGAGATATAATGTTAA	47429
QY	2161	ATTAATCTCACTCACTATAGTATGTCGTTTATATATATATATAATAAATTTATTAATTC	2220
DB	47428	ATTAATCTCACTCACTATAGTATGTCGTTTATATATATATATAATAAATTTATTAATTC	47369
QY	2221	CATAAATTTATGTTATGATTAATTTGGAGACCTATATCATATATATAACAGAAATTTGTA	2280
DB	47368	CATAAATTTATGTTATGATTAATTTGGAGACCTATATCATATATATAACAGAAATTTGTA	47309
QY	2281	TGAATGACAGACTTATACCTTATGTTAGTAGCTTTTATCTCTGTCATAATGTTATTTGA	2340
DB	47308	TGAATGACAGACTTATACCTTATGTTAGTAGCTTTTATCTCTGTCATAATGTTATTTGA	47249
QY	2341	CTTTGTCCTTAAATTTTTTTTAAATATATTTGTTGTTATTTCTTTTCAGCGGTTTA	2400
DB	47248	CTTTGTCCTTAAATTTTTTTTAAATATATTTGTTGTTATTTCTTTTCAGCGGTTTA	47189
QY	2401	TGTCACGCTTGTCAAATTTGGTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAAT	2460
DB	47188	TGTCACGCTTGTCAAATTTGGTACACAGCTGATTTTATTTAGACATGCTACGCTTTTAAT	47129
QY	2461	TATTTCTTTTCCATTTTCTATTTTATAATCTGATATATAATTTAGTCTACTTTTA	2520
DB	47128	TATTTCTTTTCCATTTTCTATTTTATAATCTGATATATAATTTAGTCTACTTTTA	47069
QY	2521	CTTCTCTCTAGTGTGAATTTT 2541	
DB	47068	CTTCTCTCTAGTGTGAATTTT 47048	
RESULT 10			
AP480875			
LOCUS	AP480875	1669 bp	DNA linear
DEFINITION	Homo sapiens prostrate-specific membrane antigen-like enhancer sequence.		PRI 14-MAR-2002
ACCESSION	AP480875		
VERSION	AP480875.1	GI:19423473	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1669)		
AUTHORS	O'Keefe,D.S., Bacich,D.J. and Heston,W.D.W.		
TITLE	Expression Profile of Prostate-Specific Membrane Antigen (PSMA) Tissues, Prostate Cancer and Tumor Associated Vasculature		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1669)		
AUTHORS	O'Keefe,D.S., Bacich,M.E., Bacich,D.J. and Heston,W.D.W.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-FEB-2002) Cancer Biology, Cleveland Clinic		

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 246865)
Waterston,K.H.
Direct Submission
Submitted (09-Jul-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, S
MO 63108, USA
On Sep 24, 2000 this sequence version replaced gi:9186850.

COMMENT

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0163214
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-primer ET; 100% of reads
Assembly: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 213368 bases at least Q40
Consensus quality: 223527 bases at least Q30
Consensus quality: 229418 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 241663; sum-of-contigs
Quality coverage: 4.83 in Q20 bases; agarose-fp
Quality coverage: 3.37 in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 'contigs' (the true order of the pieces
* is not known) and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved

*	1	12900:	contig	of 12900 bp	in length
*	12901	13000:	gap	of unknown length	
*	13001	23006:	contig	of 10006 bp	in length
*	23007	23106:	gap	of unknown length	
*	23107	35032:	contig	of 11926 bp	in length
*	35033	35132:	gap	of unknown length	
*	35133	50532:	contig	of 15400 bp	in length
*	50533	50632:	gap	of unknown length	
*	50633	64756:	contig	of 14124 bp	in length
*	64757	64856:	gap	of unknown length	
*	64857	80978:	contig	of 16122 bp	in length
*	80979	81078:	gap	of unknown length	
*	81079	10159:	contig	of 20081 bp	in length
*	101160	101259:	gap	of unknown length	
*	101260	102865:	contig	of 1606 bp	in length
*	102866	102965:	gap	of unknown length	
*	102966	104593:	contig	of 1628 bp	in length
*	104594	104893:	gap	of unknown length	
*	104694	106231:	contig	of 1538 bp	in length
*	106232	106331:	gap	of unknown length	
*	106332	107693:	contig	of 1362 bp	in length
*	107694	107793:	gap	of unknown length	
*	107794	109155:	contig	of 1362 bp	in length
*	109156	109255:	gap	of unknown length	
*	109256	110481:	contig	of 1226 bp	in length
*	110482	110581:	gap	of unknown length	
*	110582	111981:	contig	of 1300 bp	in length
*	111882	111981:	gap	of unknown length	
*	111982	11324:	contig	of 1543 bp	in length
*	11325	113624:	gap	of unknown length	
*	113625	114983:	contig	of 1359 bp	in length
*	114984	115083:	gap	of unknown length	
*	115084	116145:	contig	of 1062 bp	in length
*	116146	116245:	gap	of unknown length	
*	116246	117625:	contig	of 1380 bp	in length
*	117626	117725:	gap	of unknown length	
*	117726	120009:	contig	of 2284 bp	in length

	source	FEATURES
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
Location/Qualifiers
1. .246865
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

120010	120109: gap of unknown length
120110	122469: contig of 2360 bp in length
120111	122569: gap of unknown length
120170	122569: gap of unknown length
122470	124001: contig of 1432 bp in length
122570	124001: contig of 1432 bp in length
124002	124101: gap of unknown length
124102	126352: contig of 2251 bp in length
126353	126352: gap of unknown length
126453	128477: contig of 2025 bp in length
128478	128577: gap of unknown length
128578	130466: contig of 1889 bp in length
130467	130566: gap of unknown length
130567	132503: contig of 1939 bp in length
132506	132605: gap of unknown length
132606	134148: contig of 1543 bp in length
134149	134248: gap of unknown length
134249	136448: contig of 2200 bp in length
136449	136548: gap of unknown length
136549	138253: contig of 1707 bp in length
138256	138355: gap of unknown length
138356	140533: contig of 2180 bp in length
140536	140633: gap of unknown length
140636	142139: contig of 1564 bp in length
142200	142299: gap of unknown length
142300	144906: contig of 2607 bp in length
144907	145006: gap of unknown length
145007	147213: contig of 2207 bp in length
147214	147313: gap of unknown length
147314	148878: contig of 1565 bp in length
148879	148978: gap of unknown length
148979	150906: contig of 1928 bp in length
150907	151006: gap of unknown length
151007	153354: contig of 2348 bp in length
153355	153454: gap of unknown length
153455	155071: contig of 1617 bp in length
155072	155171: gap of unknown length
155172	159010: contig of 3839 bp in length
159011	159110: gap of unknown length
159111	162070: contig of 2960 bp in length
162071	162470: gap of unknown length
162471	164152: contig of 2282 bp in length
164153	164552: gap of unknown length
164553	1658294: contig of 3742 bp in length
1658295	168394: gap of unknown length
168395	171385: contig of 2991 bp in length
171386	171485: gap of unknown length
171486	175016: contig of 3531 bp in length
175017	175116: gap of unknown length
175117	178766: contig of 3650 bp in length
178767	178866: gap of unknown length
178867	182838: contig of 3972 bp in length
182839	182938: gap of unknown length
182939	186148: contig of 3210 bp in length
186149	186248: gap of unknown length
186249	190168: contig of 3920 bp in length
190169	190268: gap of unknown length
190269	193524: contig of 3256 bp in length
193525	193624: gap of unknown length
193625	202260: contig of 8536 bp in length
202261	202260: gap of unknown length
202262	210133: contig of 7873 bp in length
210134	210233: gap of unknown length
210234	218141: contig of 7908 bp in length
218142	218241: gap of unknown length
218242	228213: contig of 9972 bp in length
228214	228313: gap of unknown length
228314	237496: contig of 9183 bp in length
237497	237596: gap of unknown length
237597	246865: contig of 9269 bp in length

Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 Travans, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

COMMENT

Submitted (16-NOV-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIPR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L28659

Center clone name: 36_C_3

* NOTE: This record contains 86 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 649: contig of 649 bp in length
 650
 749: gap of 100 bp
 750
 1448: contig of 699 bp in length
 1449
 1548: gap of 100 bp
 1549
 2264: contig of 716 bp in length
 2265
 2364: gap of 100 bp
 2365
 3078: contig of 714 bp in length
 3079
 3178: gap of 100 bp
 3179
 3878: contig of 700 bp in length
 3879
 3978: gap of 100 bp
 3979
 4673: contig of 695 bp in length
 4674
 4773: gap of 100 bp
 4774
 5453: contig of 680 bp in length
 5454
 5553: gap of 100 bp
 5554
 6262: contig of 709 bp in length
 6263
 6362: gap of 100 bp
 6363
 7068: contig of 706 bp in length
 7069
 7168: gap of 100 bp
 7169
 7864: contig of 696 bp in length
 7865
 7964: gap of 100 bp
 7965
 8669: contig of 705 bp in length
 8670
 8769: gap of 100 bp
 8770
 9472: contig of 703 bp in length
 9473
 9572: gap of 100 bp
 9573
 10268: contig of 696 bp in length
 10269
 10368: gap of 100 bp
 10369
 11070: contig of 702 bp in length
 11071
 11170: gap of 100 bp
 11171
 11881: contig of 711 bp in length
 11882
 11981: gap of 100 bp
 11982
 12689: contig of 708 bp in length
 12690
 12789: gap of 100 bp
 12790
 13499: contig of 710 bp in length
 13500
 13599: gap of 100 bp
 13600
 14312: contig of 713 bp in length
 14313
 14412: gap of 100 bp
 14413
 15124: contig of 712 bp in length
 15125
 15224: gap of 100 bp
 15225
 15928: contig of 704 bp in length
 15929
 16028: gap of 100 bp
 16029
 16699: contig of 671 bp in length
 16700
 16799: gap of 100 bp
 16800
 17493: contig of 694 bp in length
 17494
 17593: gap of 100 bp

18292: contig of 699 bp in length
 18293
 18392: gap of 100 bp
 18393
 19090: contig of 698 bp in length
 19091
 19190: gap of 100 bp
 19191
 1901: contig of 711 bp in length
 1902
 20001: gap of 100 bp
 20002
 20710: contig of 709 bp in length
 20711
 20810: gap of 100 bp
 20811
 21507: contig of 697 bp in length
 21508
 21607: gap of 100 bp
 21608
 22314: contig of 707 bp in length
 22315
 22414: gap of 100 bp
 22415
 23129: contig of 715 bp in length
 23130
 23229: gap of 100 bp
 23230
 23937: contig of 708 bp in length
 23938
 24037: gap of 100 bp
 24038
 24744: contig of 707 bp in length
 24745
 24844: gap of 100 bp
 24845
 25546: contig of 702 bp in length
 25547
 25647: gap of 100 bp
 25648
 26342: contig of 696 bp in length
 26343
 26442: gap of 100 bp
 26443
 27153: contig of 711 bp in length
 27154
 27253: gap of 100 bp
 27254
 27955: contig of 702 bp in length
 27956
 28055: gap of 100 bp
 28056
 28759: contig of 704 bp in length
 28760
 28859: gap of 100 bp
 28860
 29569: contig of 710 bp in length
 29570
 29669: gap of 100 bp
 29670
 30372: contig of 703 bp in length
 30373
 30472: gap of 100 bp
 30473
 31159: contig of 687 bp in length
 31160
 31259: gap of 100 bp
 31260
 31973: contig of 714 bp in length
 31974
 32073: gap of 100 bp
 32074
 32786: contig of 713 bp in length
 32787
 32886: gap of 100 bp
 32887
 33593: contig of 707 bp in length
 33594
 33693: gap of 100 bp
 33694
 34394: contig of 701 bp in length
 34395
 34494: gap of 100 bp
 34495
 35195: contig of 701 bp in length
 35196
 35295: gap of 100 bp
 35296
 35989: contig of 694 bp in length
 35990
 36089: gap of 100 bp
 36090
 36785: contig of 696 bp in length
 36786
 36885: gap of 100 bp
 36886
 37595: contig of 710 bp in length
 37596
 37695: gap of 100 bp
 37696
 38404: contig of 709 bp in length
 38405
 38504: gap of 100 bp
 38505
 39215: contig of 711 bp in length
 39216
 39315: gap of 100 bp
 39316
 40030: contig of 715 bp in length
 40031
 40130: gap of 100 bp
 40131
 40843: contig of 713 bp in length
 40844
 40943: gap of 100 bp
 40944
 41659: contig of 716 bp in length
 41660
 41759: gap of 100 bp
 41760
 42455: contig of 696 bp in length
 42456
 42555: gap of 100 bp
 42556
 43259: contig of 704 bp in length
 43260
 43359: gap of 100 bp
 43360
 44060: contig of 701 bp in length
 44061
 44160: gap of 100 bp
 44161
 44874: contig of 714 bp in length
 44875
 44974: gap of 100 bp
 44975
 45693: contig of 719 bp in length
 45694
 45793: gap of 100 bp
 45794
 46498: contig of 705 bp in length
 46499
 46598: gap of 100 bp
 46599
 47292: contig of 694 bp in length

[illegible]

Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., MacDonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrum,J., Meneus,I., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-NOV-2002) Whitehead Institute/MIT Center for Genom
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L28635
Center clone name: 432_G_8

* NOTE: This record contains 51 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 719: contig of 719 bp in length
720 819: gap of 100 bp
820 1514: contig of 695 bp in length
1515 1614: gap of 100 bp
1615 2305: contig of 691 bp in length
2306 2405: gap of 100 bp
2406 3095: contig of 690 bp in length
3096 3195: gap of 100 bp
3196 3859: contig of 664 bp in length
3860 4647: contig of 688 bp in length
4648 4747: gap of 100 bp
4748 5453: contig of 706 bp in length
5454 5553: gap of 100 bp
5554 6260: contig of 707 bp in length
6261 7065: contig of 705 bp in length
7066 7165: gap of 100 bp
7166 7866: contig of 701 bp in length
7867 7966: gap of 100 bp
7967 8653: contig of 687 bp in length
8654 8754: gap of 100 bp
8754 9466: contig of 713 bp in length
9467 9566: gap of 100 bp
9567 10252: contig of 686 bp in length
10253 10352: gap of 100 bp
10353 11062: contig of 710 bp in length
11063 11162: gap of 100 bp
11163 11859: contig of 697 bp in length
11860 11959: gap of 100 bp
11960 12593: contig of 634 bp in length
12594 12693: gap of 100 bp
12694 13394: contig of 701 bp in length
13395 14149: gap of 100 bp
13495 14169: contig of 675 bp in length

QY	1884	ATCTCTAAATTTGAATATATTGAGATCAGATAATATGTGGTTTGTAGACACACTAATCC-TTTG	1942
Db	48843	ATCTCTAAATTTGAACATATTTGAGATCAGATAATATGTGGTTTGTAGACACACTAATCCTTTTG	48784
QY	1943	ACAATTTGTTGAGCCTTCCTTTTGGAAACCTAAATATGTGCTCAATTTTATAGACGTTCTGTG	2002
Db	48783	ACAATTTGTTGAGCCTTCCTTTTGGAAACCTNAIANGTCTCAATNNATNAATATTTGTGTG	48724
QY	2003	TTTCTTTTGGGAAAAACATATATTGATGGTTGTTTGGTTTAAATATTTTGTATTTGTACAT	2062
Db	48723	TTTCTTTTGGGAAAAACATATATTGATGGTTGTTTGGTTTAAATATTTTGTATTTGTACAT	48664
QY	2063	TAGTTTGAGTTTGCTTATTATTATTGGCTGAAATCTCAATATCCCTTAATGCTCTCTCAT	2122
Db	48663	TAGTTTGAGTTTGCTTATTATTATTGGCTGAAATCTCCATATCCCTTAATGCTCTCTCAT	48604
QY	2123	TTTGTCTGCTTCCTTTTATTAAATTTAGAGATAAAATGTTTAAATATCTCACCTCACTATAGTG	2182
Db	48603	TTTGTCTGCTTCCTTTTATTAAATTTAGAGATAAAATGTTTAAATATCTCACCTCACTATAGTA	48544
QY	2183	ATGCTGCTTTTATACATATATATAAAATTTTAAATCTCATAAATTTATGTTATGTATATA	2242
Db	48543	ATGCTGCTTTTATACATATATATAAAATTTTAAATCTCATAAATTTTATGTTATGTATATA	48484
QY	2243	TTTGGAGACCTTATATCATATATAAACC-----AGAAATGTTTGATGAAGATGACACACTTA	2296
Db	48483	TTTGGAGACCTTATTAACATATACAAACAAATTTAGAAATAGTTTGAAGAAATGACACACTTA	48424
QY	2297	TACTTATGATAGTAGCCCTTTTTTATCTCGTCATAAATGTTATTTGACTTTTGTCTCTAAATTT	2356
Db	48423	TACTTATGATAGTAGCCCTTTTTTATCTCGTCATAAATGTTATTTGACTTTTGTCTCTAAATTT	48364
QY	2357	TTTTTT-AAATTAATTTGTTTGGTATTTCTTTTTCAGCGGGTTTATGTCACGTTGTGCA	2415
Db	48363	TTTTTTAAATTAATTTGTTTGGTATTTCTTTTTCAGTAGGTTTATGTCACGTTGTGCA	48304
QY	2416	ATTGGTACACAGCTGATTTTATTAGACATGCTACGCTTTTAAATTTATCTTTTTCCTAT	2475
Db	48303	ATTGGTACACAGCTGATTTTATTAGACATGCTACGCTTTTAAATTTATCTTTTTCCTAT	48244
QY	2476	TTTCAATTTTTTAAATCTCATATACAATATTAGCTCACTTTTACCTTCCTAGTGTG	2535
Db	48243	TTTCAATTTTTTAAATCTCATATACAATATTAGCTCACTTTTACCTTCCTAGTGTG	48184
QY	2536	AAATTTT 2541	
Db	48183	AAATTTT 48178	
RESULT 14			
AC136711			
LOCUS		40131 bp	DNA
DEFINITION		linear	HTG 07-NOV-2002
		40131 bp	map 15, LOW-PASS
			SEQUENCE SAMPLING.
AC136711			
VERSION		GI:24756997	
KEYWORDS		HTG; HTGS PHASE0	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE			
AUTHORS		1 (bases 1 to 40131)	
TITLE		Birren, B., Nusbaum, C. and Lander, E.	
JOURNAL		Homo sapiens chromosome 15, clone RP13-432G8	
REFERENCE		2 (bases 1 to 40131)	
AUTHORS		Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,	
		Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,	
		Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,	
		Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,	
		Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,	
		Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 17:08:16 ; Search time 935.344 Seconds
(without alignments)

10695.346 Million cell updates/sec

Title: US-09-914-651A-1

Perfect score: 335

Sequence: 1 aattatttttctttaacc.....tgcaacctcgctcccg9g 335

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estlin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_estc.*

9: gb_est1.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_tod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	93.2	27.8	414	13	BX504814
C 2	92.8	27.7	684	14	CD642149
C 3	89.2	26.6	839	14	CD520912
C 4	89	26.6	1034	13	BX431378

C 5	88.8	26.5	486	14	CD238898
C 6	88	26.3	867	13	BU561446
C 7	87.2	26.0	1479	11	BC032458
C 8	86.8	25.9	509	10	BE144987
C 9	86.6	25.9	609	9	AV760466
C 10	86	25.7	600	9	AV759518
C 11	85.8	25.6	398	9	AL696113
C 12	85.8	25.6	727	14	CD000795
C 13	85.6	25.5	598	28	AQ053689
C 14	84.6	25.3	660	9	AL704150
C 15	84.4	25.2	585	13	BQ778512
C 16	84	25.1	223	9	AL697921
C 17	84	25.1	405	14	CD237959
C 18	84	25.1	562	9	AL704374
C 19	83.8	25.0	683	9	AV733431
C 20	83.8	25.0	745	13	BU564170
C 21	83.6	25.0	679	29	AG084405
C 22	83.2	24.8	933	13	BQ897609
C 23	83	24.8	822	12	BG575169
C 24	82.8	24.7	213	9	AL712392
C 25	82.8	24.7	400	9	AV760391
C 26	82.8	24.7	469	9	AL042753
C 27	82.8	24.7	689	29	AG123600
C 28	82.8	24.7	1104	11	BC031345
C 29	82.8	24.7	1723	11	BC039095
C 30	82.4	24.6	378	13	BU954284
C 31	82.4	24.6	722	14	CD639592
C 32	82.2	24.5	328	13	BX645511
C 33	82	24.5	485	28	AQ134713
C 34	82	24.5	928	13	BX452229
C 35	81.8	24.4	329	13	BX645928
C 36	81.8	24.4	362	9	AL712440
C 37	81.6	24.4	667	29	AG150532
C 38	81.6	24.4	884	10	BF184192
C 39	81.4	24.3	485	28	AQ488128
C 40	81.4	24.3	939	14	CD172331
C 41	81.2	24.2	210	13	BX476439
C 42	81.2	24.2	376	14	CD107262
C 43	81	24.2	537	9	AL157651
C 44	81	24.2	728	29	AG013357
C 45	80.6	24.1	212	10	AW772536

ALIGNMENTS

RESULT 1
BX504814/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX504814 414 bp mRNA linear EST 04-SEP-2003
DKFZp686J05182.r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DKFZp686J05182.5', mRNA sequence.

BX504814

BX504814

BX504814.1 GI:32032215

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 414)

Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,

Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and

Wiemann, S.

EST (Poustka, A., Albert, R., Moosmayer, P., Schupp, I.,

Wellenreuther, R., et al.)

Unpublished (2003)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by DKFZ (German Cancer Research Center,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.
No s1 sequence available.
This clone (DKFZp686J05182) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J05182"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplEx2; Site_1: SfiII; Site_2: SfiII; cDNA-collection"

FEATURES

source
1. .414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J05182"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplEx2; Site_1: SfiII; Site_2: SfiII; cDNA-collection"

ORIGIN

Query Match 27.8%; Score 93.2; DB 13; Length 414;
Best Local Similarity 65.7%; Pred. No. 1.2e-07;
Matches 134; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 132 TTTTATTGTTGTTAAAGACTTTATGTTGTCATATGATAAATTTATGTCATAGAAC 191
Db 402 TTTTNTTT 343
QY 192 AATTTTATGTCATATGACATTTATGTTATATGCTATACAACTAGATTTTTTTTG 251
Db 342 TTTTNTTT 283
QY 252 TTGTTTTTGACGAGTCTTACTGTGACCCAGGTGGAGTGTATGGATGCTCAGC 311
Db 282 TTTTCTGAGACAGAGTCTTCTGTGACCCAGGTGGAGTGTATGGCGGATCTCAGC 223
QY 312 TCATGCAACCTCCGCTCCCGG 335
Db 222 TCATGCAACCTCCGCTCCCGG 199

RESULT 2

CD642149/c
LOCUS
DEFINITION
AGENCOURT_14537030 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:30417982 5', mRNA sequence.
CD642149
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 684)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM220 row: d column: 23
High quality sequence stop: 214.
Location/Qualifiers
1. .684
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES

source
1. .684
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:30417982"
/tissue_type="Pooled"
/lab_host="DH10B (T1 phase-resistant)"
/clone_lib="NIH_MGC_191"
/note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
Site_2: SfiI (ggcgctcgcc); Library is oligo-dT primed
and directionally cloned. PBM - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCCCATTTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCACATG-TT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 27.7%; Score 92.8; DB 14; Length 684;
Best Local Similarity 61.7%; Pred. No. 1.2e-07;
Matches 148; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 93 TTATTGGTTGTAATATGACTATTTTACGTCATATAACAATTTTATTGTTGTTAAATGA 152
Db 242 TTTTNTTT 183
QY 153 CTTTATTGTTGTCATATGATAAATTTATGTCATAGACAAATTTTATTGCTTGATATAT 212
Db 182 TTTTNTTT 123
QY 213 GACTTTATTGTTATGCTATACAACTAGATTTTTCGTTGTTTGACCGAGCTTAC 272
Db 122 TTTTNTTT 63
QY 273 TCTGTACCCAGCTGGAGTGAATGGCATGTCTCAGCTCACTGCAACTCGGCTCC 332
Db 62 TCTGTACCCAGCTGGAGTGAATGGCATGTCTCAGCTCACTGCAACTCGGCTCC 3

RESULT 3

CD520912/c
LOCUS
DEFINITION
AGENCOURT_14356130 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:30413219 5', mRNA sequence.
CD520912
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 839)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM207 row: n column: 12
High quality sequence stop: 305.
Location/Qualifiers
1. .839
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES

source
1. .839
/organism="Homo sapiens"
/mol_type="mRNA"

[illegible]

RESULT 6	867 bp	mRNA	linear	EST 16-SEP-2002
BUS611446/c				
LOCUS				
DEFINITION	AGENCOURT_10278336 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6592490			
	5', mRNA sequence.			
ACCESSION	BUS611446			
VERSION	BUS611446.1			
KEYWORDS	GI:22911742			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 867)			
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .			
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)			
TITLE	Unpublished (1999)			
JOURNAL	Contact: Robert Strausberg, Ph.D.			
COMMENT	Email: cgapbs@email.nih.gov			
	Tissue Procurement: CLONTECH			
	cDNA Library Preparation: CLONTECH Laboratories, Inc.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Agencourt Bioscience Corporation			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLCM2815	row: m	column: 02	
	High quality sequence	stop: 343.		

```

FEATURES             Location/Qualifiers
     1..867
         organism="Homo sapiens"
         molecule_type="mRNA"
         db_xref="taxon:9606"
         clone="IMAGE:6592490"
         lab_host="DH10B (T1 phage-resistant)"
         clone_lib="NIH MGC 82"
         notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgccctcgccg); Site 2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGACCGAGGGCGGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```

```

ORIGIN
Query Match          26.3%; Score 88; DB 13; Length 867;
Best Local Similarity 59.7%; Pred. No. 8.3e-07;
Matches 148; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 88 TTCTCTTTATTGGGTGTAAATGACTATTTTACGTCATATACAAATTTTATTTCTGTTA 147
Db 328 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 269

```

QY	148	AATGACTTTATTCTTGGTCATATGATAAAATTTATGTCAAGAACAAATTTATTATGCTGA	207
Db	268	TTT	209
QY	208	TATATGACCTTTATTTGTTTATATGGCTATACAACATAGATTTTTTTTGTTGTTTTTGACCGAGT	267
Db	208	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTATTTATTTATTTTTCAGACGGA	149
QY	268	CTTACTCTGTCACCCAGGCTGGAGTGTAAATGGCATGGTCTCAGCTCACTGC AACCTCCGC	327
Db	148	GTCCTCGTGTCACCAGGCTGGAGTGAGCACCACTCAGCTTACTACAACTCCGC	89
QY	328	CTCCCCGG 335	
Db	88	CTCCCCGG 81	

RESULT 7

BC032458/c	BC032458	1479 bp	mRNA	linear	HTC 04-MAR-2003
LOCUS	Homo sapiens, clone IMAGE:5200744, mRNA.				
DEFINITION	Homo sapiens, clone IMAGE:5200744, mRNA.				
ACCESSION	BC032458				
VERSION	BC032458.1 GI:22749744				
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1479)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				

REMARK
COMMENT

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcapsb@email.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nhri.nih.gov Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskerip, B., Mastrian, S.D., McCloskey, J. C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.D., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK P.A. 54 Row: h Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES source

```
i.1479
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5200744"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_lib="NIH MGC_114"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
```

```

ORIGIN
Query Match      26.0%; Score 87.2; DB 11; Length 1479;
Best Local Similarity 64.2%; Pred. No. 9.7e-07;
Matches 131; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 132 TTTTATTTGTTGTTAAATGACCTTTATTTGTTGTCATATGATAAATTTATGTCATAGAAC 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1473 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 192 AATTTTATGCTTCATATATGACCTTTATTTGTTATATGCTATACAACTAGATTTTATTTG 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1413 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1354
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 252 TGTGTTTACCGAGCTTCTCTGTCACCCAGGCTGGAGTGTAAAGCATGCTCAGC 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1353 TTTTGTAGATGGAGTCTCACTCTGTCACCCAGGCTGGAGTGTAAAGCATGCTCAGC 1294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 312 TCAGTCAACCTCCGCTCCCGGG 335
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1293 TCAGTCAACCTCCAGCTCCCGGG 1270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
LOCUS      BE144987
DEFINITION RC2-HT0187-271099-011-c05 HT0187 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE144987
VERSION    BE144987.1 GI:8607711
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 509)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Macukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC2-HT0187-271
099-011-c05&t3=1999-10-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 503.
FEATURES
Location/Qualifiers
1..509
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0187"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match      25.9%; Score 86.8; DB 10; Length 509;
Best Local Similarity 59.2%; Pred. No. 1.7e-06;
Matches 148; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 86 AATTTCTTTATTTGGTTGAATATGACTATTTTACGTCAATACAAATTTTATTTGTTGT 145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 AAATAAATATCAGTTTATTAATCTTTTAAAAACACGCTTTAGGTTTGTTCATTTC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 146 TAAATGACTTTATTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTTGCTT 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TCTTCTACTTATCCATTTTCCACTTACTGTTTCTGCTCTTATACATATTAATATGCTTC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 206 GATATATGACTTTATTTGTTATATGCTATACAACTAGATTTTATTTGTTGTTTTCACCGA 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 TGTACTTACTTTGGGTTTATTTTGTCTATTTTGTGTTTGGTTTGTAGTTTTCAGACA 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 266 GTCTTACTCTGTCACCCAGGCTGGAGTGTAAAGCATGCTCAGCTCACTGCAACCTCC 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GTCTCATTTTGTCAACCCAGGCTGGAGTGTGATCTCAACTCACTGCAACCTCC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 326 GCCTCCCGGG 335
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 GCCTCCCGGG 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
LOCUS      AV760466/c
DEFINITION AV760466 MDS Homo sapiens cDNA clone MDSBKFL1 5', mRNA sequence.
ACCESSION AV760466
VERSION    AV760466.1 GI:10918314
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 609)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,
Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA MDS clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1..609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MDSBKFL1"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="MDS"
/note="Vector: pTriplEx2; Site_1: sfiI; Site_2: sfiIIB"

ORIGIN
Query Match      25.9%; Score 86.6; DB 9; Length 609;
Best Local Similarity 60.3%; Pred. No. 1.7e-06;
Matches 161; Conservative 0; Mismatches 104; Indels 2; Gaps 1;

QY 69 TTTTAACTACTGGTTTAAATTTCTTTATTTGTTGTTGAATATGACTATTTTACGTATATAA 128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

RESULT 10	600 bp	linear	EST 19-OCT-2000
AV759518/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

```

FEATURES
source
Location/Qualifiers
1. .600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MDSAEH04"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="MDS"
/note="vector: pYripleX2; Site 1: sf1A; Site 2: sf1B"

```

[illegible]

D β	330	T T T T A T C C C C T G A G A T G G A G T C T	271
Q γ	270	T A C T C T G T C A C C A G G C T G G A G T G T A A T G C A C T G G T C T C A G C T C A C T G C A A C C T C C G C C T	329
D δ	270	C G T C T C T C C A C C C A G G C T G G A G T G C A G T C C T T G C C T C A C T G C A A G C T C C G C C T	211
Q γ	330	C C C G G G 335 	
D δ	210	C C T G G G 205 	

[illegible]

ACCESSION	AL696113
VERSION	AL696113.1
KEYWORDS	GI:19618031
EST.	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 398)
AUTHORS	Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.)
TITLE	Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.)
COMMENT	EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.)
JOURNAL	Unpublished (1999)
CONTACT	Contact: MIPS

MTPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by DKFZ (German Cancer Research Center,
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No s1 sequence available.
 This clone (DKFZp686A04101) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg. GERMANY; Email: clone@rzpd.de

```

FEATURES
source
Berlin-Charrington9
Location/Qualifiers
1..398
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686A04101"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cdna_collection"

```

[illegible]

RESULT 12
 CD000795
 LOCUS
 DEFINITION
 AGENCOURT 13645840 NIH MGC 186 Homo sapiens cDNA clone
 IMAGE:30321777 5', mRNA sequence.
 CD000795
 VERSION
 CD000795.1 GI:30295326
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 727)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Falkovits
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDCM121 row: 1 column: 10
 High quality sequence stop: 528.
 Location/Qualifiers
 1..727
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30321777"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 186"
 /note="Organ: Pooled Skin; Vector: pDNR-LIB; Site 1: SfiI
 (ggcattatggcc); Site 2: SfiI (ggccgctggcc); Library is
 oligo-dT primed and directionally cloned. cDNA was
 prepared from a pooled samples of tissues from skin,
 meninges, duramater, pia matter and choroid plexus. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGGAGCGGCATG-dt(30)BN-3'
 (where B = A, C, G, or T). Average
 insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH MGC
 Library"

FEATURES

source

Location/Qualifiers

1..727
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30321777"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 186"
 /note="Organ: Pooled Skin; Vector: pDNR-LIB; Site 1: SfiI
 (ggcattatggcc); Site 2: SfiI (ggccgctggcc); Library is
 oligo-dT primed and directionally cloned. cDNA was
 prepared from a pooled samples of tissues from skin,
 meninges, duramater, pia matter and choroid plexus. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGGAGCGGCATG-dt(30)BN-3'
 (where B = A, C, G, or T). Average
 insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH MGC
 Library"

ORIGIN

Query Match 25.6%; Score 85.8; DB 14; Length 727;
 Best Local Similarity 59.0%; Pred. No. 2.2e-06;
 Matches 147; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
 QY 87 ATTCTTTATGGTGTAAATGACTATTATTCGTCATATACAAATTTTATGTTTGT 146
 DB 159 ATTTAATTATAGAACAGACTAAATTTTATAATTTATTTAATATATATATACATAT 218
 QY 147 AAATGACTTTATGTTGTTCATATGATTAATTTTATGTCATAGAACAAATTTTATGCTTG 206
 DB 219 ATATGTTGTATATATATACATATATATATGTTGTATATATATATATATATATATGTT 278
 QY 207 ATATATGACTTTTATGTTTATATGGCTATACAACTAGATTTTTTGTGTTTGTCCGAG 266
 DB 279 GGTGTATATATGATGTAT 338
 QY 267 TCTTACTCTGTCAACCCAGCTGGAGTGAATGGCATGTCTCAGCTCACTGCAACCTCG 326
 DB 339 TCTCACTCTGTCAACCCAGCTGGAGTGAATGGCATGTCTCAGCTCACTGCAACCTCTG 398

QY 327 CCTCCCGG 335
 DB 399 CCTCCCGG 407
 AQQ53689 598 bp DNA linear GSS 20-APR-1999
 RPC111-51010 TK RPC1-11 Homo sapiens genomic clone RPC1-11-51010,
 genomic survey sequence.
 AQQ53689
 ACCESSION
 AQQ53689.1 GI:3348612
 KEYWORDS
 GSS.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 598)
 Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
 Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and
 Venter, J.C.
 Use of BAC End Sequences for Sequence-Ready Map Building (1998)
 Unpublished (1998)
 COMMENT
 Other GSSs: RPC111-51010.TJ
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadam@tigr.org
 Clones are derived from the human BAC library RPC1-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
 Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Class: BAC ends.
 Location/Qualifiers
 1..598
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7519545"
 /db_xref="taxon:9606"
 /clone="RPC1-11-51010"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPC1-11"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
 RPC111 Human Male BAC Library"

FEATURES

source

1..598
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7519545"
 /db_xref="taxon:9606"
 /clone="RPC1-11-51010"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPC1-11"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
 RPC111 Human Male BAC Library"

ORIGIN

Query Match 25.6%; Score 85.6; DB 28; Length 598;
 Best Local Similarity 60.2%; Pred. No. 2.6e-06;
 Matches 142; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
 QY 100 TTGTAATGACTATTATTCAGTCATATACAAATTTTATGTTTGTAAATGACTTTAT 159
 DB 286 TTAATAACATTTAATTTACATACAAATTTTAAATATTTTAAAAAGTTATATATATATA 345
 QY 160 GTTGTCTCATGATAAATTTTATGTCATAGAACAAATTTTATGTTGCTTGATATGACTTTA 219
 DB 346 TACATA 405
 QY 220 TTGTTATATGGCTATACAACTAGATTTTTTTTGTGTTTTCACCGAGTCTTACTCTGTCA 279
 DB 406 TACATATATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 465
 QY 280 CCCAGCTGGAGTGTATGGCATGTCTCAGCTCACTCAACCTCCGCTCCCGG 335
 DB 466 CCCAGCTGGAGTGTATGGCATGTCTCAGCTCACTCAACCTCCGCTCCCGG 521

RESULT 14
 AL704150/c
 LOCUS
 DEFINITION DKFZp686D0529_r1 686 (synonym: hicc3) Homo sapiens cDNA clone
 AL704150
 ACCESSION
 VERSION DKFZp686D0529_5', mRNA sequence.
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 660)
 AUTHORS Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
 EST (Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., et al.)
 JOURNAL Unpublished (2003)
 COMMENT Contact: MIPS

MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No s1 sequence available.
 This clone (DKFZp686D0529) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1..660
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686D0529"
 /lab_host="adult"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hicc3)"
 /note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
 cDNA-collection"

ORIGIN
 Query Match 25.3%; Score 84.6; DB 9; Length 660;
 Best Local Similarity 68.4%; Pred. No. 3.8e-06;
 Matches 117; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 165 TCATATGATATTTTATGCTATGACAAATTTTATGCTGTATATATGCTTATTTGTT 224
 Db 427 TCATGCGCATATGATTTTATATATATATATATGCTGTGTGTGTGTGTGTGTATAT 368
 QY 225 ATATGCTATACAACTAGATTTTGTGTTTGTGACCGAGCTTACTCTGTCAACCCAG 284
 Db 367 ATATACATTACAAATCTATTTTCTTTTAAACAGAGCTCTACTCTGTGCGCCGG 308
 QY 285 GCTGGAGTGAATGGCATGCTTCTAGCTACTGCAACCTCCGCTCCCGG 335
 Db 307 GCTGGAGTGAATGGCATGCTTCTAGCTACTGCAACCTCCGCTCCCGG 257

RESULT 15
 BQ778512
 LOCUS
 DEFINITION BQ778512 585 bp mRNA linear EST 26-JUL-2002
 mRNA sequence.
 ACCESSION BQ778512
 VERSION BQ778512.1 GI:21986984
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 585)
 AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

TITLE
 JOURNAL
 COMMENT

Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R.,
 Williams, T., Jackson, Y., and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other ESTs: il32e06.y1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 Seq primer: -400P from Gibco
 High quality sequence stop: 447.

FEATURES
 source

1..585
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6031595"
 /tissue type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
 NotI; Site_2: XhoI; cDNA made by oligo-dT priming. ~1kb. 5'
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 25.2%; Score 84.4; DB 13; Length 585;
 Best Local Similarity 66.5%; Pred. No. 4.4e-06;
 Matches 121; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 154 TTTATTTGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATTTGCTTGATATG 213
 Db 18 TTTCATTTTATTAATTCGACACACATTTGATTGACAAATATTCATTCAGGATATGTC 77
 QY 214 ACTTTATGTTATATGGCTATACACTAGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 273
 Db 78 CTTCATATCCATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTGCT 137
 QY 274 CTGTACCCAGGCTGAGTGTATGTCATGCTCAGCTCACGCAACCTCCGCTCCCG 333
 Db 138 CTTCACCCAGGCTGAGTGTATGTCATGCTCAGCTCACGCAACCTCTGCCCCCG 197
 QY 334 GG 335
 Db 198 GG 199

Search completed: May 21, 2004, 23:03:19
 Job time : 939.344 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 20:49:36 ; Search time 151.193 Seconds
(without alignments)
10069.828 Million cell updates/sec

Title: US-09-914-651A-1

Perfect score: 335

Sequence: 1 aattatttttcttaacc.....tgcacacctccgctcccggg 335

Scoring table: IDENTITY NUC

Gapop 10.0 ; Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	312	93.1	335	US-10-431-791-3	Sequence 3, Appli
2	262	78.2	454	US-10-431-791-1	Sequence 1, Appli
3	246	73.4	257	US-10-431-791-18	Sequence 18, Appl
4	94	28.1	13444	US-09-764-877-3660	Sequence 3660, Ap
5	94	28.1	13444	US-10-242-515-3660	Sequence 3660, Ap
6	89.8	26.8	53332	US-09-801-861-3	Sequence 3, Appli
7	89.8	26.8	53332	US-10-224-562-3	Sequence 3, Appli
8	86.4	25.8	40941	US-10-087-192-1426	Sequence 1426, Ap
9	85.4	25.5	567	US-10-027-632-138568	Sequence 138568,
10	85.4	25.5	567	US-10-027-632-138569	Sequence 138569,
11	85.4	25.5	567	US-10-027-632-138570	Sequence 138570,
12	85.4	25.5	567	US-10-027-632-138568	Sequence 138568,
13	85.4	25.5	567	US-10-027-632-138569	Sequence 138569,
14	85.4	25.5	567	US-10-027-632-138570	Sequence 138570,

15	85.2	25.4	2185	16	US-10-074-024-821	Sequence 821, App
16	85	25.4	237326	15	US-10-301-844-2	Sequence 2, Appli
17	85	25.4	378361	10	US-09-901-136-3	Sequence 3, Appli
18	84.8	25.3	22197	13	US-10-087-192-1156	Sequence 1156, Ap
19	84.6	25.3	133632	13	US-10-087-192-1810	Sequence 1810, Ap
20	83.8	25.0	72604	15	US-10-162-497-7	Sequence 7, Appli
21	83.6	25.0	18998	9	US-09-764-869-1952	Sequence 1952, Ap
22	83.6	25.0	18998	15	US-10-091-504-1952	Sequence 1952, Ap
23	83.6	25.0	18998	16	US-10-227-577-1952	Sequence 1789, Ap
24	83.6	25.0	108316	16	US-10-232-798-1789	Sequence 2143, Ap
25	83.6	25.0	108317	15	US-10-017-161-2143	Sequence 118828,
26	83.4	24.9	1108	13	US-10-027-632-118828	Sequence 214024,
27	83.4	24.9	1108	16	US-10-027-632-118828	Sequence 214024,
28	83.2	24.8	650	13	US-10-027-632-116384	Sequence 116384,
29	83.2	24.8	650	16	US-10-027-632-116384	Sequence 116384,
30	82.8	24.7	3015	13	US-10-027-632-116383	Sequence 116383,
31	82.8	24.7	3015	13	US-10-027-632-116383	Sequence 116384,
32	82.8	24.7	3015	13	US-10-027-632-116384	Sequence 116385,
33	82.8	24.7	3015	16	US-10-027-632-116385	Sequence 116383,
34	82.8	24.7	3015	16	US-10-027-632-116384	Sequence 116384,
35	82.8	24.7	3015	16	US-10-027-632-116385	Sequence 116385,
36	82.6	24.7	762	13	US-10-027-632-163521	Sequence 163521,
37	82.6	24.7	762	16	US-10-027-632-163521	Sequence 163521,
38	82.4	24.6	1119	13	US-10-027-632-117192	Sequence 117192,
39	82.4	24.6	1119	13	US-10-027-632-117192	Sequence 117193,
40	82.4	24.6	1119	16	US-10-027-632-117192	Sequence 117192,
41	82.4	24.6	1119	16	US-10-027-632-117192	Sequence 117193,
42	82.4	24.6	147309	9	US-09-742-312-3	Sequence 3, Appli
43	82.4	24.6	147309	15	US-10-436-185-3	Sequence 3, Appli
44	82.2	24.5	687	13	US-10-027-632-27766	Sequence 27766, A
45	82.2	24.5	687	16	US-10-027-632-27766	Sequence 27766, A

ALIGNMENTS

RESULT 1
US-10-431-791-3
; Sequence 3, Application US/10431791
; Publication No. US20030235874A1
; GENERAL INFORMATION:
; APPLICANT: Kao, Chinghai
; APPLICANT: Lee, Sang-Jin
; APPLICANT: Kim, Hong-Sup
; APPLICANT: Lee, KangRyul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; FILE REFERENCE: 1857-ARTI.0222US
; CURRENT APPLICATION NUMBER: US/10/431,791
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-431-791-3

Query Match 93.1%; Score 312; DB 16; Length 335;
Best Local Similarity 99.4%; Pred. No. 5.9e-53;
Matches 334; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 AATTATTTTCTTAACTTCACTTCAACTTCAACTTCAAGGAAACAGTTCCTTGTCTGTGTT 60
Db 1 AATTATTTTCTTAACTTCACTTCAACTTCAAGGAAACAGTTCCTTGTCTGTGTT 60
QY 61 GTGGAATTTTAACTTCACTTCACTTCAACTTCAAGGAAACAGTTCCTTGTCTGTGTT 120

Db 61 GTGGAATAATTTAAACTACTGGTTTAAATTTCTTTATTGGTTTGTAAATATGACTATTTTACG 120
Qy 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATT-GTTTGTCAATATGATAAATTTT 179
Db 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTGTCAATATGATAAATTTT 180
Qy 180 ATGTCATAGAACAAATTTTATTGTTTGTATATATGACTTTATTGTTTGTATATGCTATACAAAC 239
Db 181 ATGTCATAGAACAAATTTTATTGTTTGTATATATGACTTTATTGTTTGTATATGCTATACAAAC 239
Qy 240 TAGATTTTGTGTTTGTACCGAGTCTTACTCTGTCAACCGCTGGAGTGAATGG 299
Db 240 TAGATTTTGTGTTTGTACCGAGTCTTACTCTGTCAACCGCTGGAGTGAATGG 299
Qy 300 CATGCTCTCAGTCTCACTCAACTCCGCTCCCGGG 335
Db 300 CATGCTCTCAGTCTCACTCAACTCCGCTCCCGGG 335

RESULT 2
US-10-431-791-1
; Sequence 1, Application US/10431791
; Publication No. US20030235874A1
; GENERAL INFORMATION:
; APPLICANT: Kao, Chinghai
; APPLICANT: Lee, Sang-Jin
; APPLICANT: Kim, Hong-Sup
; APPLICANT: Lee, Kangryul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; FILE REFERENCE: 1857-ARTI.0222US
; CURRENT APPLICATION NUMBER: US/10/431,791
; CURRENT FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-431-791-1

Query Match 78.2%; Score 262; DB 16; Length 454;
Best Local Similarity 100.0%; Pred.No. 6.1e-43; Indels 0; Gaps 0;
Matches 262; Conservative 0; Mismatches 0;
Qy 1 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAACCAAGTTGGCTTGGACTCTGTTT 60
Db 193 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAACCAAGTTGGCTTGGACTCTGTTT 252
Qy 61 GTGGAATAATTTAACTACTGGTTTAAATTTCTTTATTGGTTTGTAAATATGACTATTTTACG 120
Db 253 GTGGAATAATTTAACTACTGGTTTAAATTTCTTTATTGGTTTGTAAATATGACTATTTTACG 312
Qy 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTGTCAATATGATAAATTTA 180
Db 313 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTGTCAATATGATAAATTTA 372
Qy 181 TGTCAATAGAACAAATTTTATTGTTTGTATATGACTTTATTGTTTGTATATGCTATACAACT 240
Db 373 TGTCAATAGAACAAATTTTATTGTTTGTATATGACTTTATTGTTTGTATATGCTATACAACT 432
Qy 241 AGATTTTGTGTTTGTGAC 262
Db 433 AGATTTTGTGTTTGTGAC 454
RESULT 3
US-10-431-791-18

; Sequence 18, Application US/10431791
; Publication No. US20030235874A1
; GENERAL INFORMATION:
; APPLICANT: Kao, Chinghai
; APPLICANT: Lee, Sang-Jin
; APPLICANT: Kim, Hong-Sup
; APPLICANT: Lee, Kangryul
; APPLICANT: Yu, Rong
; TITLE OF INVENTION: Prostate-Specific Chimeric Enhancers and
; FILE REFERENCE: 1857-ARTI.0222US
; CURRENT APPLICATION NUMBER: US/10/431,791
; CURRENT FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/378,920
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-431-791-18

Query Match 73.4%; Score 246; DB 16; Length 257;
Best Local Similarity 99.6%; Pred.No. 7.7e-40;
Matches 257; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAACCAAGTTGGCTTGGACTCTGTTT 60
Db 1 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAACCAAGTTGGCTTGGACTCTGTTT 60
Qy 61 GTGGAATAATTTAACTACTGGTTTAAATTTCTTTATTGGTTTGTAAATATGACTATTTTACG 120
Db 61 GTGGAATAATTTAACTACTGGTTTAAATTTCTTTATTGGTTTGTAAATATGACTATTTTACG 120
Qy 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTGTCAATATGATAAATTTA 180
Db 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTGTCAATATGATAAATTTA 180
Qy 181 TGTCAATAGAACAAATTTTATTGTTTGTATATGACTTTATTGTTTGTATATGCTATACAACT 240
Db 181 TGTCAATAGAACAAATTTTATTGTTTGTATATGACTTTATTGTTTGTATATGCTATACAACT 239
Qy 241 AGATTTTGTGTTTGT 258
Db 240 AGATTTTGTGTTTGT 257

RESULT 4
US-09-764-877-3660/c
; Sequence 3660, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3660
; LENGTH: 13444
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3660

Query Match 28.1%; Score 94; DB 9; Length 13444;
Best Local Similarity 65.0%; Pred.No. 6.6e-09;
Matches 139; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 122 CATATAACAATTTTATGTTTAAATGACTTTATTTGTTTGTCTATGATGATAATTTTAT 181
Db 4821 CATCTTAAATTTTCCCTTCITTTTAAATATGTTGTAATCTATTTATTTTGGTCTTCAT 4762
QY 182 GTCATAGAACAAATTTTATGCTTGATATGACTTTTATTTGTTATATGCTATACACTA 241
Db 4761 GTTAAATACAGAAATCTGTTCTAGGTAGATTTTTCAGTGTCAATTTGGGATTTTCACTTT 4702
QY 242 GAATTTTGTGTTTGTGACCGAGTCTTACTCTGTCAACCGAGGCTGGAGTGTAAATGSCA 301
Db 4701 TTTTATTTTGTGTTTGTGACCGAGTCTTCTGTCTGCGCCAGCGGAGTGTGCGG 4642
QY 302 TGGTCTCAGTCACTGCAACTCCGCTCCCGG 335
Db 4641 TGATCTCGGCTCACTGCAAGCTCCGCTCCCGG 4608

RESULT 5

US-10-242-515-3660/c
; Sequence 3660, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; PRIORITY FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3660
; LENGTH: 13444
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-3660

Query Match 28.1%; Score 94; DB 16; Length 13444;
Best Local Similarity 65.0%; Pred. No. 6.6e-09;
Matches 139; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 122 CATATAACAATTTTATGTTTAAATGACTTTATTTGTTTGTCTATGATGATAATTTTAT 181
Db 4821 CATCTTAAATTTTCCCTTCITTTTAAATATGTTGTAATCTATTTATTTTGGTCTTCAT 4762
QY 182 GTCATAGAACAAATTTTATGCTTGATATGACTTTTATTTGTTATATGCTATACACTA 241
Db 4761 GTTAAATACAGAAATCTGTTCTAGGTAGATTTTTCAGTGTCAATTTGGGATTTTCACTTT 4702
QY 242 GAATTTTGTGTTTGTGACCGAGTCTTACTCTGTCAACCGAGGCTGGAGTGTAAATGSCA 301
Db 4701 TTTTATTTTGTGTTTGTGACCGAGTCTTCTGTCTGCGCCAGCGGAGTGTGCGG 4642
QY 302 TGGTCTCAGTCACTGCAACCTCCGCTCCCGG 335

Db 4641 TGATCTCGGCTCACTGCAAGCTCCGCTCCCGG 4608

RESULT 6

US-09-801-861-3
; Sequence 3, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match 26.8%; Score 89.8; DB 9; Length 53332;
Best Local Similarity 59.8%; Pred. No. 7.6e-08;
Matches 168; Conservative 0; Mismatches 112; Indels 1; Gaps 1;
QY 54 TCTGTTTGTGGAATAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTAATGACTA 113
Db 31467 TATATATTTATATATATTTTATATATATTTATATATATATATATATATATTTATA 31526
QY 114 TTTTACGTCATATAACAATTTTATTTGTTTAAATGACTTTTATTTGTTGTCATATGAT 173
Db 31527 TATATTTATATATTTTATATATATTTATATATATTTATATATATTTATATATTTA 31586
QY 174 AATTTTATGTCATAGAACAAATTTTATTTGCTTCATATATGACTTTATTTGTTATGCTA 233
Db 31587 TATATTTATATATTTTATATATTTTATATATTTATATATTTA-TATATATTTTATATA 31645
QY 234 TACAACCTAGATTTTGTGTTGTTTGTGACCGAGTCTTACTCTGTCAACCGGCTGGAGTG 293
Db 31646 TATATATATGATTTTGTGTTTGTGAGATGAGTCTCACTCTATTGCCAGGCTGGAGTG 31705
QY 294 TAATGGCATGCTCTCAGTCACTGCAACCTCCGCTCCCGG 334
Db 31706 CAGTGGCAGCATCTCAGTCACTGCAACCTCCACCTCCCGG 31746

RESULT 7

US-10-224-562-3
; Sequence 3, Application US/10224562
; Publication No. US20030022229A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-562-3

Query Match 26.8%; Score 89.8; DB 15; Length 53332;
Best Local Similarity 59.8%; Pred. No. 7.6e-08;
Matches 168; Conservative 0; Mismatches 112; Indels 1; Gaps 1;

Qy		54	TCTGTTTGCGAAAATTTTTAAACTACGTGGTTCATTATTCGGTGGTAATAATGACTA	113
Dβ		31467	TATATATTTTATATATATTTTATATATPATTTATATATATATATATATATATTTATATTTATA	31526
Qy		114	TTTTACGCATATAACAAGTTTTTATGTTGTTTAAAGACTTTATGTTTGGCATATGAT	173
Dβ		31527	TATATTTATATATTTATATATATTTATATATATTTATATATATTTATATATATTTA	31586
Qy		174	AATTTTATGTCATAGAACAATTTTATTTGCTTGATATATGACTTTATTGTTTATATGGCTA	233
Dβ		31587	TATATTATATATATTTTATATATTTATATATTTTATATATTA-TATATATATTTTATATATA	31645
Qy		234	TACAACTAGATATTTTGTGTTGTTTTGACGGAGCCTTACTCTGTCACCACCGAGCTGGAGTG	293
Dβ		31646	TATATATATGTAATTTTTTTTTTTTGAGATGGAGCTCACTCTATTGCCCCAGCTGGAGTG	31705
Qy		294	TAATGGCATGGTCTCTCACCTCACTGC AACCTCCGCCTCCC GG	334
Dβ		31706	CAGTGGACAGCATCTTCAGCTCACTGC AACTCCA CACTCCCC AG	31746

RESULT 8
US-10-087-192-1426
; Sequence 1426, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

```

; TITLE OF INVENTION:  CANCEER
;
; FILE REFERENCE:  529452000122
;
; CURRENT APPLICATION NUMBER:  US/10/087,192
;
; CURRENT FILING DATE:  2002-03-01
;
; PRIOR APPLICATION NUMBER:  US 09/747,377
;
; PRIOR FILING DATE:  2000-12-22
;
; PRIOR APPLICATION NUMBER:  US 09/798,586
;
; PRIOR FILING DATE:  2001-03-02
;
; NUMBER OF SEQ ID NOS:  2059
;
; SOFTWARE:  fastSEQ for Windows Version 4.0
;
; SEQ ID NO 1426
;
; LENGTH:  40491
;

```

	Query Match	25.8%;	Score 86.4;	DB 13;	Length 40491;
	Best Local Similarity	73.0%;	Pred. No. 3.3e-07;		
	Matches 111;	Conservative 0;	Mismatches 41;	Indels 0;	Gaps 0;
QY	184	CATAGAACAAATTTTATTGCTTGATATATACACTTTATTGTATATCGCTATACAACACTAGA	243		
Db	11828	CATAATTTTT	11887		
QY	244	TTTTTTTTGTCGTTTTTCACCGAGCTTACTCTGTCCACCAGGCTGGAGTGTAATGCCATG	303		
Db	11888	TTCGTTTTTTAATTGACACGGAGTCTTCCTGTGCACACAGGCTGGAGTGCATGCCATG	11947		
QY	304	GTCTCAGCTCACTGCAACCTCCGCTCCCGGG	335		
Db	11948	AFTCTGGCTCACTGCAAGCTGCTGCTCCCGGG	11979		

```

RESULT 9
US-10-027-632-138568
; Sequence 138568, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

```

```

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138568
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138568

Query Match          25.5%;      Score 85.4
Best Local Similarity 65.4%;      Pred. No. 1
Matches 125;          Conservative 0; Mismatch

QY      145  TTAATGACTTTATTGTTTGTTCATATGATGATA
Db      298  TAAAAAAAATTCATTGGATGGCATTTATATAGG
QY      205  TGATATATGACTTTATCTTTATATATGGCTATT
Db      358  TGTTTTTATTATTATTATTATTATTATTATTATT
QY      265  AGTCTTACTCTGTCAACCAGGCTGGAGTGTT
Db      418  AGTCTTGTCTGTCTCCAGGCTGGAGTGTC
QY      325  GCGCTCCCGG 335
Db      478  CACCTCCGG 488

```

```

RESULT 10
US-10-027-632-138569
; Sequence 138569, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138569
; LENGTH: 567
; TYPE: DNA

```

```

; ORGANISM: Human
US-10-027-632-138569

Query Match      25.5%; Score 85.4; DB 13; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACTTTATGTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTGGATGGCAATTTATAGCATCTTAGGTGAACAAAGAAATTTGTT 357
QY 205 TGATATATGACTTTATGTTATATGCTATACACTAGATTTTGTGTTTGGACCG 264
Db 358 TGTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTGGATGG 417
QY 265 AGTCTTACTCTGACCCAGGCTGGAGTGTATGCGCATGTTCTCAGCTCACTGCAACCTC 324
Db 418 AGCTTGTCTGTCTCCAGGCTGGAGTGGCAATCTCAGCTCACTGCAAACTC 477
QY 325 CGCCTCCGGG 335
Db 478 CACCTTCGGG 488

RESULT 11
US-10-027-632-138570
; Sequence 138570, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138570
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138570

Query Match      25.5%; Score 85.4; DB 13; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACTTTATGTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTGGATGGCAATTTATAGCATCTTAGGTGAACAAAGAAATTTGTT 357
QY 205 TGATATATGACTTTATGTTATATGCTATACACTAGATTTTGTGTTTGGACCG 264
Db 358 TGTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTGGATGG 417
QY 265 AGTCTTACTCTGACCCAGGCTGGAGTGTATGCGCATGTTCTCAGCTCACTGCAACCTC 324
Db 418 AGCTTGTCTGTCTCCAGGCTGGAGTGGCAATCTCAGCTCACTGCAAACTC 477
QY 325 CGCCTCCGGG 335
Db 478 CACCTTCGGG 488

RESULT 12
US-10-027-632-138568
; Sequence 138568, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138568
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138568

Query Match      25.5%; Score 85.4; DB 16; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 145 TTAATGACTTTATGTTGTCATATGATAAATTTATGTCATAGAACAAATTTTATTGCT 204
Db 298 TAAAAAAATTCATTGGATGGCAATTTATAGCATCTTAGGTGAACAAAGAAATTTGTT 357
QY 205 TGATATATGACTTTATGTTATATGCTATACACTAGATTTTGTGTTTGGACCG 264
Db 358 TGTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTGGATGG 417
QY 265 AGTCTTACTCTGACCCAGGCTGGAGTGTATGCGCATGTTCTCAGCTCACTGCAACCTC 324
Db 418 AGCTTGTCTGTCTCCAGGCTGGAGTGGCAATCTCAGCTCACTGCAAACTC 477
QY 325 CGCCTCCGGG 335
Db 478 CACCTTCGGG 488

RESULT 13
US-10-027-632-138569
; Sequence 138569, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```

Query Match      25.5%; Score 85.4; DB 16; Length 567;
Best Local Similarity 65.4%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

145  TTAATGACTTTATTGTTGTCATATGATAATTTTATGTCATAGAAACAATTTTATTGCT 204
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298  TAAAAAAAATTCATTGGATGGCATTTATAGCATCTTAGGTGAACAAAAGAAATTTGGT 357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
358  TGTTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTGATGG 417
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265  AGTCCTTACTCTGTCACCCAGGCTGGAGTGTAAATGGCATGGTCTCAGCTCACTGCAACCTC 324
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418  AGCTTTGCTCTGTCCTCCAGGCTGGAGTGGCAGTGGCAAACTTCAGCTCACTGCAAACTC 477
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

325  CGCCTCCCGGG 335
      ||||| |||||
478  CACCTTCGGG 488

RESULT 15
US-10-074-024-821
; Sequence 821, Application US/10074024
; Publication No. US20030232975A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC001C1
; CURRENT APPLICATION NUMBER: US/10/074,024
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 879
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 821
; LENGTH: 2185
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-024-821

```

```

Query Match      25.4%; Score 85.2; DB 16; Length 2185;
Best Local Similarity 66.1%; Pred. No. 1.9e-07;
Matches 123; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 150 TGACTTTATGTTTGCATATGATATTTTATGTCATAGACAAATTTTATTGCTTCGATA 209
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 365 TAAATTTTATTTTATATATATACACATATATACATGATATACATATGTTGTGTG 424
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 210 TATGACTTTTATGTTTATATGCGCTATACAACTAGATTTTTTTTGTGTTTTTGACCGAGTCT 269
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 425 TGTGTGTGTGTGTGTGTGTATATATATATATTTTTTTTTTTTGTGACACAGAGTCT 484
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 270 TACTCTGTACCCAGCGCTGGAGTGTAATGGCATGGTCTCAGCTCAGTCGAACCTCCGCT 329
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 485 TACTCTGTATCCAGCGCTGGAGTACAGTGAATCTCAGCTCAGTCGAACCTTCGCT 544
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 330 CCGGG 335
      |||||
Db 545 CTTGG 550

```

Search completed: May 22, 2004, 02:28:08
Job time : 154.193 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	89.8	26.8	53332	4	US-09-801-861-3	Sequence 3, Appli
C 2	83.8	25.0	72604	4	US-09-268-932-7	Sequence 7, Appli
C 3	83.8	25.0	72604	4	US-09-657-474-7	Sequence 7, Appli
4	83.2	24.8	28720	4	US-09-341-587-7	Sequence 7, Appli
C 5	81.8	24.4	62804	4	US-09-800-960-3	Sequence 3, Appli
C 6	81.8	24.4	62804	4	US-10-096-960-3	Sequence 3, Appli
C 7	81.6	24.4	6799	4	US-09-620-312D-299	Sequence 299, App
8	79.8	23.8	28001	4	US-09-819-933-3	Sequence 3, Appli
C 9	79.8	23.8	28001	4	US-10-193-295-3	Sequence 3, Appli
10	79	23.6	36159	4	US-09-749-558-3	Sequence 3, Appli
C 11	78.6	23.5	11811	3	US-09-078-294-7	Sequence 7, Appli
C 12	77.8	23.2	246240	2	US-08-724-394A-20	Sequence 7, Appli
13	77.8	23.2	246240	2	US-08-724-394A-21	Sequence 20, Appl
14	77.8	23.2	246240	2	US-08-724-394A-22	Sequence 21, Appl
C 15	77.4	23.1	29629	4	US-09-729-995-3	Sequence 22, Appl
C 16	77.4	23.1	29629	4	US-10-135-689-3	Sequence 3, Appli
C 17	77.2	23.0	49312	4	US-09-671-317-485	Sequence 3, Appli
C 18	77.2	23.0	392000	4	US-10-027-983-11	Sequence 485, App
C 19	77	23.0	12597	4	US-09-705-239-12	Sequence 11, Appl
20	76.8	22.9	3701	4	US-09-220-132-57	Sequence 52, Appl
C 21	76.8	22.9	84495	4	US-09-797-906-3	Sequence 3, Appli
C 22	76.6	22.9	38844	4	US-09-734-675-3	Sequence 3, Appli
C 23	76.2	22.7	4428	4	US-09-023-655-1109	Sequence 3, Appli
C 24	76.2	22.7	6038	3	US-09-305-639-4	Sequence 1109, Ap
C 25	76.2	22.7	6038	3	US-09-525-160B-2	Sequence 4, Appli
C 26	76.2	22.7	7622	3	US-09-305-639-1	Sequence 2, Appli
C 27	76.2	22.7	7622	4	US-09-535-160B-1	Sequence 1, Appli

APPLICANT: Chen, H.
APPLICANT: Freimer, N.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
FILE REFERENCE: 7853-138
CURRENT APPLICATION NUMBER: US/09/268,992
CURRENT FILING DATE: 1999-03-16
EARLIER APPLICATION NUMBER: 09/236,134
EARLIER FILING DATE: 1999-01-22
EARLIER APPLICATION NUMBER: 60/106,056
EARLIER FILING DATE: 1998-10-28
EARLIER APPLICATION NUMBER: 60/088,312
EARLIER FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 60/078,044
EARLIER FILING DATE: 1998-03-16
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 72604
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified base
LOCATION: all n positions
OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match 25.0%; Score 83.8; DB 4; Length 72604;
Best Local Similarity 57.8%; Pred. No. 6.9e-11;
Matches 145; Conservative 2; Mismatches 104; Indels 0; Gaps 0;

QY 85 TAATTTCTTTTATGGTTGTAATATGACTATTTTACGTCATATAAACAATTTTATGTTG 144
Db 64262 TTATTTAAATTTTTCGAATATTTTAAACGTTTCATAAAATGATTCGAGTGTATTAGTCA 64203

QY 145 TTAATGACTTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCT 204
Db 64202 TAATMTACTTAATCAATTCCTCTCTGTTTACACATTAAAGACTATGTTTAAACATATATAT 64143

QY 205 TGAATATGACTTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCT 264
Db 64142 ATAAT 64083

QY 265 AGCTTACTCTGTCACCCAGGCTGGAGTGTAAATGGCATGCTCTCAGCTCACTGCAACCTC 324
Db 64082 AGCTTCTCTGTCGCCAGGCTGGAGTGGAGTGGCGGATCTCTGCTCACTGCAAGTTC 64023

QY 325 CGCTCCCGG 335
Db 64022 TGCCTCCCGG 64012

RESULT 4
US-09-341-587-7
Sequence 7, Application US/09341587
Patent No. 6346606
GENERAL INFORMATION:
APPLICANT: Mollenhauer, Jan
TITLE OF INVENTION: Protein Containing an SRCR Domain
FILE REFERENCE: 4121-108
CURRENT APPLICATION NUMBER: US/09/341,587
CURRENT FILING DATE: 1999-08-31
EARLIER APPLICATION NUMBER: PCT/DE98/00096
EARLIER FILING DATE: 1998-01-09
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 28720
TYPE: DNA
ORGANISM: Homo sapiens
US-09-341-587-7

Query Match 24.8%; Score 83.2; DB 4; Length 28720;
Best Local Similarity 58.5%; Pred. No. 7.8e-11;
Matches 145; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 88 TTCTTTTATGGTTGTAATATGACTATTTTACGTCATATAAACAATTTTATGTTGTTGA 147
Db 19249 TTCTTTCT 19308

QY 148 AATGACTTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCTTGA 207
Db 19309 TCCTTTCT 19368

QY 208 TATATGACTTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCTTGA 267
Db 19369 TCT 19428

APPLICANT: Chen, H.
APPLICANT: Freimer, N.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
FILE REFERENCE: 7853-138
CURRENT APPLICATION NUMBER: US/09/268,992
CURRENT FILING DATE: 1999-03-16
EARLIER APPLICATION NUMBER: 09/236,134
EARLIER FILING DATE: 1999-01-22
EARLIER APPLICATION NUMBER: 60/106,056
EARLIER FILING DATE: 1998-10-28
EARLIER APPLICATION NUMBER: 60/088,312
EARLIER FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 60/078,044
EARLIER FILING DATE: 1998-03-16
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 72604
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified base
LOCATION: all n positions
OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match 25.0%; Score 83.8; DB 4; Length 72604;
Best Local Similarity 57.8%; Pred. No. 6.9e-11;
Matches 145; Conservative 2; Mismatches 104; Indels 0; Gaps 0;

QY 85 TAATTTCTTTTATGGTTGTAATATGACTATTTTACGTCATATAAACAATTTTATGTTG 144
Db 64262 TTATTTAAATTTTTCGAATATTTTAAACGTTTCATAAAATGATTCGAGTGTATTAGTCA 64203

QY 145 TTAATGACTTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCT 204
Db 64202 TAATMTACTTAATCAATTCCTCTCTGTTTACACATTAAAGACTATGTTTAAACATATATAT 64143

QY 205 TGAATATGACTTTTATGTTGTCATATGATAATTTTATGTCATAGAACAAATTTTATGCT 264
Db 64142 ATAAT 64083

QY 265 AGCTTACTCTGTCACCCAGGCTGGAGTGTAAATGGCATGCTCTCAGCTCACTGCAACCTC 324
Db 64082 AGCTTCTCTGTCGCCAGGCTGGAGTGGAGTGGCGGATCTCTGCTCACTGCAAGTTC 64023

QY 325 CGCTCCCGG 335
Db 64022 TGCCTCCCGG 64012

RESULT 3
US-09-657-474-7/c
Sequence 7, Application US/09657474
Patent No. 6399762
GENERAL INFORMATION:
APPLICANT: Chen, H.
APPLICANT: Freimer, N.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
FILE REFERENCE: 7853-138
CURRENT APPLICATION NUMBER: US/09/657,474
CURRENT FILING DATE: 2000-09-07
EARLIER APPLICATION NUMBER: 09/268,992
EARLIER FILING DATE: 1999-03-16
EARLIER APPLICATION NUMBER: 09/236,134
EARLIER FILING DATE: 1999-01-22
EARLIER APPLICATION NUMBER: 60/106,056
EARLIER FILING DATE: 1998-10-28
EARLIER APPLICATION NUMBER: 60/088,312
EARLIER FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/078,044

QY 268 CTTACTCTGTACCCAGGCTGAGTGTATATGCGCTCTCAGCTCAGTGCAGCTCCG 327
|||
Db 19429 CTCGCTGTGTACCCAGGCTGGAATGTAGTGGCAGGCTCTCGGCTCACTGCACCTCTGC 19488
|||
QY 328 CTCCCGGG 335
|||
Db 19489 CTCCCGGG 19496
|||

RESULT 5
US-09-800-960-3/c
; Sequence 3, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

Query Match 24.4%; Score 81.8; DB 4; Length 62804;
Best Local Similarity 73.8%; Pred. No. 2e-10;
Matches 104; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 195 TTTTATGCTGTATATGAGCTTATTGTTATATGCTATACAACTAGATTTTGTG 254
|||
Db 54648 TTTTATATTTATATATATATATATATATATATATATATATATATATTTT 54589
|||
QY 255 TTTTGTACCGAGTCTTACTCTGTACCCAGGCTGAGTGTATATGCGCTATACAACTAGATTTTGTG 314
|||
Db 54588 CTTGAGACAGAGTCTCGCTCTGTACCCAGGCTGAGTGTGCAATCTCAGCTCA 54529
|||
QY 315 CTGCAACTCCGCTCCCGGG 335
|||
Db 54528 CTGCAACTCCGCTCCCGGG 54508
|||

RESULT 6
US-10-096-960-3/c
; Sequence 3, Application US/10096960
; Patent No. 6664085
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001158DIV
; CURRENT APPLICATION NUMBER: US/10/096,960
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 09/800,960
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A,T,C or G
US-10-096-960-3
Query Match 24.4%; Score 81.8; DB 4; Length 62804;
Best Local Similarity 73.8%; Pred. No. 2e-10;
Matches 104; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 195 TTTTATGCTGTATATGAGCTTATTGTTATATGCTATACAACTAGATTTTGTG 254
|||
Db 54648 TTTTATATTTATATATATATATATATATATATATATATATATATTTT 54589
|||
QY 255 TTTTGTACCGAGTCTTACTCTGTACCCAGGCTGAGTGTATATGCGCTATACAACTAGATTTTGTG 314
|||
Db 54588 CTTGAGACAGAGTCTCGCTCTGTACCCAGGCTGAGTGTGCAATCTCAGCTCA 54529
|||
QY 315 CTGCAACTCCGCTCCCGGG 335
|||
Db 54528 CTGCAACTCCGCTCCCGGG 54508
|||

RESULT 7

US-09-620-312D-299/c
; Sequence 299, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 299
; LENGTH: 6799
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (249)..(2405)
US-09-620-312D-299

Query Match 24.4%; Score 81.6; DB 4; Length 6799;
Best Local Similarity 83.0%; Pred. No. 1.4e-10;
Matches 93; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 224 TATATGCTATACAACTAGATTTTGTGTTTGTGACCGAGTCTTACTCTGTACCCA 283
|||
Db 3407 TATATGCATACATTTCTATTTTTTTTTTTTGTGAGACAGAGTCTCTCTGTACCCA 3348
|||
QY 284 GGCTGGAGTGTATGCGATGCTCTCAGCTCACTGCACCTCCGCTCCCGGG 335
|||

Mon May 24 08:12:33 2004

Db 3347 GGCTGGAGTGAATGGCATGATCTCGGCTCACTGCAACCTCGCTCCAGG 3296

RESULT 8

US-09-819-993-3
; Sequence 3, Application US/09819993
; Patent No. 6436692
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001195
; CURRENT APPLICATION NUMBER: US/09/819,993
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(28001)
; OTHER INFORMATION: n = A,T,C or G
US-09-819-993-3

Query Match 23.8%; Score 79.8; DB 4; Length 28001;
Best Local Similarity 71.4%; Pred. No. 5e-10;
Matches 105; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 189 AACAAATTTTATGCTTGATATATGACTTTTATTTATGCTATACAACTAGATTTT 248
Db 16199 ATCTATGATGAATACATGATTTGATTTGATTTGATTTTCTTTT 16258
QY 249 TTGTTGTTTTTGACCGAGTCTTACTCTGTCAACCGAGTGGAGTGAATGGCATGCTC 308
Db 16259 TTTTCTGTCTGACCGAGTCTTCTGTCTGTCAACCGAGTGGAGTGAATGGCATGCTC 16318
QY 309 AGCTCACTGCAACTCGCTCCCGG 335
Db 16319 TGCTCACTGCAAGCTCTCTCTCCCGG 16345

RESULT 9

US-10-193-295-3
; Sequence 3, Application US/10193295
; Patent No. 6620608
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001195DIV
; CURRENT APPLICATION NUMBER: US/10/193,295
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 08/819,993
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(28001)
; OTHER INFORMATION: n = A,T,C or G
US-10-193-295-3

Query Match 23.8%; Score 79.8; DB 4; Length 28001;
Best Local Similarity 71.4%; Pred. No. 5e-10;
Matches 105; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 189 AACAAATTTTATGCTTGATATATGACTTTTATTTATGCTATACAACTAGATTTT 248
Db 16199 ATCTATGATGAATACATGATTTGATTTGATTTGATTTTCTTTT 16258
QY 249 TTGTTGTTTTTGACCGAGTCTTACTCTGTCAACCGAGTGGAGTGAATGGCATGCTC 308
Db 16259 TTTTCTGTCTGACCGAGTCTTCTGTCTGTCAACCGAGTGGAGTGAATGGCATGCTC 16318
QY 309 AGCTCACTGCAACTCGCTCCCGG 335
Db 16319 TGCTCACTGCAAGCTCTCTCTCCCGG 16345

RESULT 10

US-09-749-588-3
; Sequence 3, Application US/09749588
; Patent No. 6423521
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001068
; CURRENT APPLICATION NUMBER: US/09/749,588
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36159
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36159)
; OTHER INFORMATION: n = A,T,C or G
US-09-749-588-3

Query Match 23.6%; Score 79; DB 4; Length 36159;
Best Local Similarity 76.4%; Pred. No. 8.1e-10;
Matches 97; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 209 ATATGACTTTTATGCTTGATATGCTATACAACTAGATTTTGTGTTGACCGAGT 268
Db 11441 ATATCACTTTTGAGGGTTACTTCTTTTCTTTTCTTTTGTGAGACAGAGTC 11500
QY 269 TTACTCTGTCAACCGAGTGGAGTGAATGGCATGCTCAGCTCACTGCAACTCCGCC 328
Db 11501 TTACTCTGTCAACCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 11560
QY 329 TCCCGG 335
Db 11561 TCCTGGG 11567

RESULT 11

US-09-078-294-7/c
; Sequence 7, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 11811
; TYPE: DNA
; ORGANISM: BAC-F2 contig 3


```
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match      23.2%; Score 77.8; DB 2; Length 246240;
Best Local Similarity 59.7%; Pred. No. 2.3e-09;
Matches 151; Conservative 0; Mismatches 97; Indels 5; Gaps 1;

QY 83 TTTAAATCTTTTATGTTGTTAATATGACTATTTTACGTCATATAACAATTTTATTTGTT 142
Db 179076 TCTAAATTTGTTAATGCTCTCTTTTAAATGCAAAAGACACATCACTTTACCTTCTTCA 179135

QY 143 TCTTAAATGACTTTATTTGTTGTCATATGATATTAATTTATGTCATAGAACAAATTTTATTTG 202
Db 179136 TTACCAAGTCTTTGGAATTTCTCTGATCTCTAATTTCTTTGTTCAATTTATTCATATAT 179195

QY 203 CTGTGATATAGACTTTATTTGTTATATGCTGCTATACAACTAGATTTTGTGTTGTTGAC 262
Db 179196 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 179250

QY 263 CGAGTCTTACTCTGTCAACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 322
Db 179251 AGAGTCTCACACTGTCAACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 179310

QY 323 TCCGCTCCCGGG 335
Db 179311 TCTGCTCTCTGAG 179323

RESULT 14
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

Query Match      23.2%; Score 77.8; DB 2; Length 246240;
Best Local Similarity 59.7%; Pred. No. 2.3e-09;
Matches 151; Conservative 0; Mismatches 97; Indels 5; Gaps 1;

QY 83 TTTAAATCTTTTATGTTGTTAATATGACTATTTTACGTCATATAACAATTTTATTTGTT 142
Db 179076 TCTAAATTTGTTAATGCTCTCTTTTAAATGCAAAAGACACATCACTTTACCTTCTTCA 179135

QY 143 TCTTAAATGACTTTATTTGTTGTCATATGATATTAATTTATGTCATAGAACAAATTTTATTTG 202
Db 179136 TTACCAAGTCTTTGGAATTTCTCTGATCTCTAATTTCTTTGTTCAATTTATTCATATAT 179195

QY 203 CTGTGATATAGACTTTATTTGTTATATGCTGCTATACAACTAGATTTTGTGTTGTTGAC 262
Db 179196 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 179250

QY 263 CGAGTCTTACTCTGTCAACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 322
Db 179251 AGAGTCTCACACTGTCAACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCTCACTGCAACC 179310

QY 323 TCCGCTCCCGGG 335
Db 179311 TCTGCTCTCTGAG 179323

RESULT 15
US-09-729-995-3/c
; Sequence 3, Application US/09729995
; Patent No. 6436206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 29629
; TYPE: DNA
; ORGANISM: Human
; US-09-729-995-3

Query Match      23.1%; Score 77.4; DB 4; Length 29629;
Best Local Similarity 71.3%; Pred. No. 1.9e-09;
Matches 102; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 193 ATTTTATTCCTTGATATATAGACTTTTATTTGTTATATGCTGCTATACAACTAGATTTTGT 252
Db 17380 ATTTTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 17321

QY 253 TGTGTTTGGACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGAATGCGCATGGTCTCAGCT 312
Db 17320 TTTTGGAGACAGAGTCTGCTGTCGCCAGGCTGGAGTGAATGCGCATGGTCTCAGCT 17261

QY 313 CACTGCAACCTCCGCTCCCGGG 335
Db 17260 CACTGGAACCTCTGCTCCCGGG 17238

Search completed: May 21, 2004, 23:07:07
Job time : 28.8588 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:24:44 ; Search time 138.729 Seconds
(without alignments)
10258.453 Million cell updates/sec

Title: US-09-914-651A-1

Perfect score: 335

Sequence: 1 aattatttttcttacc.....tgcacacctcgctcccggg 335

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	335	100.0	335	3	Aaa94789 Human Pro
2	94	28.1	13444	4	Aal37295 Human mus
3	94	28.1	13444	7	Abx60283 cDNA enco
4	90.6	27.0	18252	4	Aak87554 Human imm
5	90.2	26.9	39110	4	Aak73087 Human imm
6	90.2	26.9	39110	4	Aak71825 Human imm
7	90.2	26.9	39110	4	Aak87555 Human imm
8	90.2	26.9	39110	7	Abz68185 Human sec
9	90.2	26.9	39110	7	Abz74663 Secreted
10	89.8	26.8	53332	6	Aal48890 Human Pft
11	87.2	26.0	140167	6	Abt10146 Human bre
12	86.4	25.8	39068	4	Aak71820 Human imm
13	86.4	25.8	39068	4	Aak85294 Human imm
14	86.4	25.8	39068	4	Aak73078 Human imm
15	86.4	25.8	39068	4	Aak87544 Human imm
16	86.4	25.8	39068	7	Abz68184 Human sec
17	86.4	25.8	39068	7	Abz74662 Secreted
18	85.8	25.6	45300	4	Aak87547 Human imm
19	85.8	25.6	45300	4	Aak73079 Human imm
20	85.8	25.6	45300	7	Abz68188 Human sec
21	85.8	25.6	45300	7	Abz74666 Secreted
22	85.2	25.4	2185	4	Aas32867 Human gen
23	85	25.4	110000	7	AAL52246_2 Continuation (3 of

c	24	85	25.4	237326	2	AAV57903
	25	84.8	25.3	227968	6	ABK83497
	26	84.2	25.1	5690	4	AAK81236
	27	84.2	25.1	5690	4	AAK81239
	28	84.2	25.1	8724	4	AAK68956
	29	83.8	25.0	72604	2	AAZ10752
c	30	83.8	25.0	72604	6	ABK43231
c	31	83.8	25.0	121600	6	ABT10748
	32	83.6	25.0	18998	4	AAK36452
	33	83.6	25.0	18998	9	AAE47146
	34	83.6	25.0	108316	9	ADC87336
	35	83.2	24.8	28720	2	AAV49655
	36	82.4	24.6	183610	7	ACF62736
c	37	82.4	24.6	183610	7	ADB20851
c	38	82.4	24.6	183610	9	ADB87940
c	39	82.4	24.6	183610	9	ADB96923
c	40	82.4	24.6	183610	9	ADB92114
c	41	82.4	24.5	21010	4	AAK89247
	42	82.2	24.5	21010	4	AAL05888
	43	82.2	24.5	21010	4	ABL98452
	44	82.2	24.5	21010	7	ABZ68148
	45	82.2	24.5	21010	7	ABZ68148

ALIGNMENTS

RESULT 1

AAA94789
ID AAA94789 standard; DNA; 335 BP.

XX AC AAA94789;

XX DT 19-JAN-2001 (first entry)

XX DE Human Prostate-Specific Membrane antigen Enhancer PSME core region.

XX KW Core region; human; prostate-specific membrane antigen; enhancer; PSME;

XX KW cancer; renal cell carcinoma; transitional cell; colonic; neuroendocrine;

XX KW malignant melanoma; pancreatic duct; breast; soft tissue; PSMA;

XX KW non-small cell lung; testicular embryonal; glioblastoma multiforme;

XX KW prostate; breast; bladder; cancer; ss.

XX OS Homo sapiens.

XX PN WO200052156-A1.

XX PD 08-SEP-2000.

XX PF 01-MAR-2000; 2000WO-AU000143.

XX PR 01-MAR-1999; 99AU-00008956.

XX PR 25-JAN-2000; 2000AU-00005268.

XX PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX PI Molloy PL, Watt F;

XX DR WPI; 2000-594182/56.

XX PT New regulatory constructs comprising intron 3 of the prostate-specific membrane antigen gene and a heterologous peptide, for directing gene expression in a prostate, bladder, breast or vascular endothelial cell.

XX PS Claim 10; Fig 11; 56pp; English.

XX CC The present sequence is the core region from human prostate-specific membrane antigen enhancer (PSME). This sequence is located in intron 3 of the PSMA gene. This sequence was used to construct a recombinant expression cassette, which is useful for the expression of proteins in vascular endothelial cells. In addition, the recombinant construct is useful in the treatment of cancer e.g. renal cell carcinoma, transitional cell carcinoma, colonic adenocarcinoma, neuroendocrine carcinoma,

CC malignant melanoma, pancreatic duct carcinoma, breast carcinoma, soft
CC tissue carcinoma, non-small cell lung carcinoma, testicular embryonal
CC carcinoma, glioblastoma multiforme, prostate, breast or bladder cancer
XX
SQ Sequence 335 BP; 84 A; 55 C; 54 G; 142 T; 0 U; 0 Other;

Query Match 100.0%; Score 335; DB 3; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.7e-54;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATTTTCTTAACTTTCACCTTCAACCTCAAGGAAACAGTGGCTTGAATCTGTTT 60
Db |||||||TTCCTTTTAACTTTCACCTTCAAGGAAACAGTGGCTTGAATCTGTTT 60

QY 61 GTGGAAATTTTAACTTTCACCTTCAACCTCAAGGAAACAGTGGCTTGAATCTGTTT 120
Db |||||||TTCCTTTTAACTTTCACCTTCAAGGAAACAGTGGCTTGAATCTGTTT 120

QY 121 TCATATAACAATTTTATTTGTTTGTAAAGACTTTATTTGTTGTCATATGATAAATTTA 180
Db |||||||TTCCTTTTAACTTTCACCTTCAAGGAAACAGTGGCTTGAATCTGTTT 180

QY 181 TGTCTAGAACAAATTTTATTTGTTTGTAAAGACTTTATTTGTTGTCATATGATAAATTTA 240
Db |||||||TTCCTTTTAACTTTCACCTTCAAGGAAACAGTGGCTTGAATCTGTTT 240

QY 241 AGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 300
Db |||||||TTCCTTTTAACTTTCACCTTCAAGGAAACAGTGGCTTGAATCTGTTT 300

QY 301 ATGGTCTCAGTCACTGCAACCTTCCGCTCCCGG 335
Db |||||||TTCCTTTTAACTTTCACCTTCAAGGAAACAGTGGCTTGAATCTGTTT 335

RESULT 2
AAL37295/C
ID AAL37295 standard; DNA; 13444 BP.
XX
AC AAL37295;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3660.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-Al.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001338.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226581P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.

28-JUN-2000; 2000US-0214886P.
30-JUN-2000; 2000US-0215135P.
07-JUL-2000; 2000US-0216647P.
11-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225213P.
14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0225266P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
18-AUG-2000; 2000US-0225759P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226868P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
25-SEP-2000; 2000US-0234998P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
13-OCT-2000; 2000US-0239937P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241212P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
20-OCT-2000; 2000US-0241826P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249246P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-02559678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.
Disclosure; SEQ ID NO 42366; 307lpp + Sequence Listing; English.
AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 18252 BP; 4805 A; 4330 C; 4265 G; 4852 T; 0 U; 0 Other;

Query Match 27.0%; Score 90.6; DB 4; Length 18252;
Best Local Similarity 60.2%; Pred. No. 2.4e-08;
Matches 150; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 87 ATTTCCTTATGTTGTTGTAATGACATGACATATTTTACGTATATACAAATTTTATGTTGTT 146
Db 786 ATATATGTTGTTGTTATATATGTTGTTATATATATATATATATATATATGTTGTTAT 845
QY 147 AATGACATTTATGTTGTTGTAATGATATATTTTATGTCATAGAACAAATTTTATGTTG 206
Db 846 ATATATACATATATATGTTGTTGTTATATATATATATATATATATGTTGTTGTTAT 905
QY 207 ATATATGACATTTATGTTTATATGCTATACAACTAGATTTTGTGTTTTCACCGAG 266
Db 906 ATATGTTGTTGTTGTTATATATATATATATATATATATATATATTTTATGACAGAG 965
QY 267 TCTTACTCTGTACCCAGGCTGAGTGATGATGCTGATGCTGCTGCTGCTGCTGCTGCTG 326
Db 966 TCTCACTCTGTACCCAGGCTGAGTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1025
QY 327 CTTCCCGGG 335
Db 1026 CTTCCCGAG 1034

RESULT 5

AAK73087
ID AAK73087 standard; DNA; 39110 BP.
XX
AC AAK73087;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27899.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209457P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.

PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
PT
PS Disclosure; SEQ ID NO 26637; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 39110 BP; 10491 A; 9099 C; 8671 G; 10849 T; 0 U; 0 Other;
Query Match 26.9%; Score 90.2; DB 4; Length 39110;

```
Best Local Similarity 61.9%; Pred. No. 2.9e-08;
Matches 143; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 105 ATATGACTATTTTACGTCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTG 164
Db 4889 ATATATACATATATACATATATACATATATATATATATATATATATATATATATAT 4830
QY 165 TCATATGATAATTTTATGTCATAGAACAAATTTTATTGTTGATATATGACTTTATTGTT 224
Db 4829 ACATATATACATATATATATATATATATATATATATATATATATATATATATATAT 4770
QY 225 ATATGGCTATACACTAGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 284
Db 4769 ATATATGATATATATATATATATATATATATATATATATATATATATATATATAT 4710
QY 285 GCTGGAGTGTAATGCACTGGTCTCAGCTCACTGCAACCTCCGCTCCCGGG 335
Db 4709 GCTGGAGTGCACTGGACACAGTCTCAGCTCACTGCACTGCACTGCTGCTCCCGAGG 4659

RESULT 7
AAK87555
ID AAK87555 standard; DNA; 39110 BP.
XX
AC AAK87555;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:42367.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
```

Db	34282	ACATATATACATATATACATATATACATATATACATATATATATATATATATATACAT	34334
Qy	225	ATATGGCTATACAACATAGATTTTTTTGTTGTTGACCGAGCTTATCTGTGACCCAG	284
Db	34342	ATATATGTATATATATATATATATTTTTTTTTTTAGACAGAGTCGCACTCTGTGACCCAG	34401
Qy	285	GCTGGAGTGTATGCGATGGTCTCAGCTCACTGCGACACCTCGCCCTCCCGG	335
Db	34402	GCTGGAGTGCAGTGGCACAGTCTCAGCTCACTGCGAGTCTCTGCTCCCGAG	34452
RESULT 8			
ABZ68185	ABZ68185 standard; DNA; 39110 BP.		
XX	AC	AC	ABZ68185;
XX	DT	DT	26-MAR-2003 (first entry)
XX	DE	DE	Human secreted protein encoding genomic DNA SEQ ID NO 1708.
XX	XX	XX	Human; secreted protein; nootropic; neuroprotective; cytostatic;
XX	XX	XX	virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW	KW	KW	vulneray; antibacterial; antiparkinsonian; antisking; antianaemic;
KW	KW	KW	antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW	KW	KW	antiinflammatory; antiallergic; antidiabetic; antituber; anticonvulsant;
KW	KW	KW	antifungal; antiparasitic; cardiac; immune disorder; infection; vaccine;
KW	KW	KW	cardiovascular disorder; neurological disease; nephrotropic;
KW	KW	KW	gene therapy; gene; ds.
OS	OS	OS	Homo sapiens.
XX	XX	XX	WO200277186-A2.
XX	XX	XX	03-OCT-2002.
XX	XX	XX	26-MAR-2002; 2002WO-US009188.
XX	XX	XX	27-MAR-2001; 2001US-0278650P.
PR	PR	PR	12-SEP-2001; 2001US-00950082.
PR	PR	PR	12-SEP-2001; 2001US-00950083.
XX	XX	XX	(HUMA-) HUMAN GENOME SCI INC.
PA	PA	PA	Rosen CA, Ruben SM;
PI	PI	PI	WPI; 2003-040583/03.
DR	DR	DR	New human secreted proteins encoded by genes contained in cDNA clones
PT	PT	PT	(e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT	PT	PT	multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT	PT	PT	West Nile fever.
XX	XX	XX	Disclosure; Page 2359-2369; 2423pp; English.
PS	PS	PS	The invention relates to novel human genes (ABZ68185-ABZ68209) and the
XX	XX	XX	encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC	CC	CC	treating or ameliorating medical conditions e.g. by protein or gene
CC	CC	CC	therapy. The genes are isolated from a range of human tissues disclosed
CC	CC	CC	in the specification. The nucleic acids, proteins, antibodies and
CC	CC	CC	(antagonists are useful in the diagnosis, treatment and prevention of:
CC	CC	CC	(a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC	CC	CC	adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC	CC	CC	lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC	CC	CC	allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC	CC	CC	diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC	CC	CC	arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC	CC	CC	myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC	CC	CC	cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC	CC	CC	bacterial, fungal and parasitic infections
XX	XX	XX	Sequence 39110 BP; 10849 A; 8671 C; 9099 G; 10491 T; 0 U; 0 Other;


```
DT 04-DEC-2002 (first entry)
DE Human breast cancer associated coding sequence SEQ ID NO: 280.
XX Human; breast specific gene; breast cancer; differential expression;
KW cytostatic; gene therapy; gene; ss.
XX Homo sapiens.
OS
XX WO200259271-A2.
PN
XX 01-AUG-2002.
PD
XX 25-JAN-2002; 2002WO-US002176.
PF
XX 25-JAN-2001; 2001US-0263757P.
PR
XX 25-APR-2001; 2001US-0286090P.
PR
XX 23-MAY-2001; 2001US-0292517P.
PR
XX (GENE-) GENE LOGIC INC.
PA
XX Orr MS, Nation M, Diggans JC, Zeng W;
PI
XX WPI; 2002-674803/72.
DR
XX
XX Diagnosing breast cancer in a patient comprises detecting the level of
PT gene expression in cell or tissue samples, where a differential gene
PT expression is indicative of breast cancer.
XX
XX Claim 1; SEQ ID NO 280; 260pp + Sequence Listing; English.
PS
XX
XX The present invention relates to methods of diagnosing breast cancer in a
CC patient, which comprise detecting the level of expression in a tissue
CC sample of two or more genes selected from those shown in ABT09867-
CC ABT1112, where a differential expression of the genes indicates breast
CC cancer. The methods are useful in diagnosing, treating, detecting the
CC progression, and in monitoring treatment of breast cancer in patients.
CC The methods are also useful as a screening tool for agents that modulate
CC the onset or progression of breast cancer. The breast cancer genes may be
CC used as diagnostic markers for the prediction or identification of the
CC malignant state of breast tissue, for confirming the type and progression
CC of cancer, and for drug screening and assays. The present sequence is a
CC coding sequence of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub.published_pct_sequences
XX
XX Sequence 140167 BP; 45038 A; 27502 C; 26800 G; 40827 T; 0 U; 0 Other;
SQ
Query Match 26.0%; Score 87.2; DB 6; Length 140167;
Best Local Similarity 69.2%; Pred. No. 1.1e-07;
Matches 119; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
Qy 164 GTCATATGATATTTTATGTATAGAACAAATTTTATGCTTGATATATGACTTTATGT 223
Db 123721 GCCAGAGATAATGCCAGATAGGTAGTATGTTTGGTATATACATATATATATA 123662
Qy 224 TATATGGCTATACACTAGATTTTTTCTGTTTGTACCGAGTCTTACTCTGTCAACCA 283
Db 123661 TATATATATATATATATATATATTTTCTTGGATGGAGTCTTCTGTCAACCA 123602
Qy 284 GGCTGGAGTGTATGGCATGGTCTCAGTCTCACTGCAACCTCGCTCCCGGG 335
Db 123601 GGCTGGAGTTCAGTGGCATGATCTCAGTCTCACTGCAAGCCGCCACCGGG 123550
RESULT 12
AAK71820/c
ID AAK71820 standard; DNA; 39068 BP.
XX
XX AAK71820;
AC
XX
XX 06-NOV-2001 (first entry)
DT
```

```
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26632.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS
XX Homo sapiens.
XX WO200157182-A2.
PN
XX 09-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US001354.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR
XX 04-FEB-2000; 2000US-0180628P.
PR
XX 24-FEB-2000; 2000US-0184664P.
PR
XX 02-MAR-2000; 2000US-0186350P.
PR
XX 16-MAR-2000; 2000US-0189874P.
PR
XX 17-MAR-2000; 2000US-0190076P.
PR
XX 18-APR-2000; 2000US-0198123P.
PR
XX 19-MAY-2000; 2000US-0205151P.
PR
XX 07-JUN-2000; 2000US-0209467P.
PR
XX 28-JUN-2000; 2000US-0214886P.
PR
XX 30-JUN-2000; 2000US-0215135P.
PR
XX 07-JUL-2000; 2000US-0216647P.
PR
XX 11-JUL-2000; 2000US-0216880P.
PR
XX 11-JUL-2000; 2000US-0217487P.
PR
XX 14-JUL-2000; 2000US-0217496P.
PR
XX 14-JUL-2000; 2000US-0218290P.
PR
XX 26-JUL-2000; 2000US-0220963P.
PR
XX 26-JUL-2000; 2000US-0220964P.
PR
XX 14-AUG-2000; 2000US-0224518P.
PR
XX 14-AUG-2000; 2000US-0224519P.
PR
XX 14-AUG-2000; 2000US-0225213P.
PR
XX 14-AUG-2000; 2000US-0225214P.
PR
XX 14-AUG-2000; 2000US-0225266P.
PR
XX 14-AUG-2000; 2000US-0225267P.
PR
XX 14-AUG-2000; 2000US-0225268P.
PR
XX 14-AUG-2000; 2000US-0225270P.
PR
XX 14-AUG-2000; 2000US-0225447P.
PR
XX 14-AUG-2000; 2000US-0225757P.
PR
XX 14-AUG-2000; 2000US-0225758P.
PR
XX 14-AUG-2000; 2000US-0225759P.
PR
XX 18-AUG-2000; 2000US-0226279P.
PR
XX 22-AUG-2000; 2000US-0226681P.
PR
XX 22-AUG-2000; 2000US-0226868P.
PR
XX 22-AUG-2000; 2000US-0227182P.
PR
XX 23-AUG-2000; 2000US-0227009P.
PR
XX 30-AUG-2000; 2000US-0228924P.
PR
XX 01-SEP-2000; 2000US-0229287P.
PR
XX 01-SEP-2000; 2000US-0229343P.
PR
XX 01-SEP-2000; 2000US-0229344P.
PR
XX 01-SEP-2000; 2000US-0229345P.
PR
XX 05-SEP-2000; 2000US-0229509P.
PR
XX 05-SEP-2000; 2000US-0229513P.
PR
XX 06-SEP-2000; 2000US-0230437P.
PR
XX 06-SEP-2000; 2000US-0230438P.
PR
XX 08-SEP-2000; 2000US-0231242P.
PR
XX 08-SEP-2000; 2000US-0231243P.
PR
XX 08-SEP-2000; 2000US-0231244P.
PR
XX 08-SEP-2000; 2000US-0231413P.
PR
XX 08-SEP-2000; 2000US-0231414P.
PR
XX 08-SEP-2000; 2000US-0232080P.
PR
XX 08-SEP-2000; 2000US-0232081P.
PR
XX 12-SEP-2000; 2000US-0231968P.
PR
XX 14-SEP-2000; 2000US-0232397P.
PR
XX 14-SEP-2000; 2000US-0232398P.
PR
XX 14-SEP-2000; 2000US-0232399P.
PR
XX 14-SEP-2000; 2000US-0232400P.
PR
XX 14-SEP-2000; 2000US-0232401P.
PR
XX 14-SEP-2000; 2000US-0233063P.
PR
XX 14-SEP-2000; 2000US-0233064P.
```

```
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250310P.
PR 01-DEC-2000; 2000US-0250319P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251475P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 26632; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 39068 BP; 10474 A; 9094 C; 8680 G; 10820 T; 0 U; 0 Other;

Query Match 25.8%; Score 86.4; DB 4; Length 39068;
Best Local Similarity 56.8%; Pred. No. 1.5e-07;
Matches 159; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 56 TGTTTGGGAAATTTAAACTACTGTTTAAATTTCTTTTATTTGTTTGTGTAATGACTATT 115
DB 32919 TATATATACATATATACATATATATATATATATATATATATATATATATATATACA 32860

QY 116 TTACGTCATATAACAATTTTATTTGTTGTTAAATGACTTTTATTTGTTGTCATATGATAA 175
DB 32859 TATATACATATATACATATATACATATATATATATATATATATATATATATATACA 32800

QY 176 TTTTATGTCATAGAACAAATTTTATTTGCTTGATATATGACTTTTATTTGTTATATGGCTATA 235
DB 32799 TACATACATATACATATATATATATATATATATATATATATATATATATATATGTA 32740

QY 236 CAACTAGATTTTGTGTTGTTTGTACCGAGTCTTACTCTGTCCACCGAGCTGGAGTGTA 295
DB 32739 TATATATTTTGTGTTTGTGTTTGTAGACAGAGTCGACTCTGTCCACCGAGCTGGAGTGCA 32680

QY 296 ATGCGATGGTCTCAGCTCACTGCAACCTCCGCTCCCGGG 335
DB 32679 GTGGCAGAGTCTCAGCTCACTGCGAGTCTCTGCTCCCGAGG 32640

RESULT 13
AAK85294/c
ID AAK85294 standard; DNA; 39068 BP.
XX
XX AAK85294;
AC
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40106.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
```

XX
OS
XX
PN
XX
XX
PD
XX
PF
XX
17-JAN-2001; 2001WO-US001354.
XX
PR
31-JAN-2000; 2000US-0179065P.
PR
04-FEB-2000; 2000US-0180628P.
PR
24-FEB-2000; 2000US-0184664P.
PR
02-MAR-2000; 2000US-0186350P.
PR
16-MAR-2000; 2000US-0189874P.
PR
17-MAR-2000; 2000US-0190076P.
PR
18-APR-2000; 2000US-0198123P.
PR
19-MAY-2000; 2000US-0205515P.
PR
07-JUN-2000; 2000US-0209467P.
PR
28-JUN-2000; 2000US-0214886P.
PR
30-JUN-2000; 2000US-0215135P.
PR
07-JUL-2000; 2000US-0216647P.
PR
07-JUL-2000; 2000US-0216880P.
PR
11-JUL-2000; 2000US-0217487P.
PR
11-JUL-2000; 2000US-0217496P.
PR
14-JUL-2000; 2000US-0218290P.
PR
26-JUL-2000; 2000US-0220963P.
PR
26-JUL-2000; 2000US-0220964P.
PR
14-AUG-2000; 2000US-0224518P.
PR
14-AUG-2000; 2000US-0224519P.
PR
14-AUG-2000; 2000US-0225213P.
PR
14-AUG-2000; 2000US-0225214P.
PR
14-AUG-2000; 2000US-0225266P.
PR
14-AUG-2000; 2000US-0225267P.
PR
14-AUG-2000; 2000US-0225268P.
PR
14-AUG-2000; 2000US-0225270P.
PR
14-AUG-2000; 2000US-0225447P.
PR
14-AUG-2000; 2000US-0225757P.
PR
14-AUG-2000; 2000US-0225758P.
PR
14-AUG-2000; 2000US-0225759P.
PR
18-AUG-2000; 2000US-0226279P.
PR
22-AUG-2000; 2000US-0226681P.
PR
22-AUG-2000; 2000US-0226868P.
PR
22-AUG-2000; 2000US-0227182P.
PR
23-AUG-2000; 2000US-0227009P.
PR
30-AUG-2000; 2000US-0228924P.
PR
01-SEP-2000; 2000US-0229287P.
PR
01-SEP-2000; 2000US-0229343P.
PR
01-SEP-2000; 2000US-0229344P.
PR
01-SEP-2000; 2000US-0229345P.
PR
05-SEP-2000; 2000US-0229509P.
PR
05-SEP-2000; 2000US-0229513P.
PR
06-SEP-2000; 2000US-0230437P.
PR
06-SEP-2000; 2000US-0230438P.
PR
08-SEP-2000; 2000US-0231242P.
PR
08-SEP-2000; 2000US-0231243P.
PR
08-SEP-2000; 2000US-0231244P.
PR
08-SEP-2000; 2000US-0231413P.
PR
08-SEP-2000; 2000US-0231414P.
PR
08-SEP-2000; 2000US-0232080P.
PR
08-SEP-2000; 2000US-0232081P.
PR
12-SEP-2000; 2000US-0231968P.
PR
14-SEP-2000; 2000US-0232397P.
PR
14-SEP-2000; 2000US-0232398P.
PR
14-SEP-2000; 2000US-0232399P.
PR
14-SEP-2000; 2000US-0232400P.
PR
14-SEP-2000; 2000US-0232401P.
PR
14-SEP-2000; 2000US-0233063P.
PR
14-SEP-2000; 2000US-0233064P.
PR
21-SEP-2000; 2000US-0233065P.
PR
21-SEP-2000; 2000US-0234223P.
PR
25-SEP-2000; 2000US-0234274P.
PR
25-SEP-2000; 2000US-0234997P.
PR
25-SEP-2000; 2000US-0234998P.

PR
26-SEP-2000; 2000US-0235484P.
PR
27-SEP-2000; 2000US-0235834P.
PR
27-SEP-2000; 2000US-0235836P.
PR
29-SEP-2000; 2000US-0236327P.
PR
29-SEP-2000; 2000US-0236367P.
PR
29-SEP-2000; 2000US-0236368P.
PR
29-SEP-2000; 2000US-0236369P.
PR
02-OCT-2000; 2000US-0236370P.
PR
02-OCT-2000; 2000US-0236802P.
PR
02-OCT-2000; 2000US-0237037P.
PR
02-OCT-2000; 2000US-0237038P.
PR
02-OCT-2000; 2000US-0237039P.
PR
13-OCT-2000; 2000US-0237040P.
PR
13-OCT-2000; 2000US-0239335P.
PR
20-OCT-2000; 2000US-0240960P.
PR
20-OCT-2000; 2000US-0241221P.
PR
20-OCT-2000; 2000US-0241785P.
PR
20-OCT-2000; 2000US-0241786P.
PR
20-OCT-2000; 2000US-0241787P.
PR
20-OCT-2000; 2000US-0241808P.
PR
20-OCT-2000; 2000US-0241809P.
PR
01-NOV-2000; 2000US-0241828P.
PR
08-NOV-2000; 2000US-0244617P.
PR
08-NOV-2000; 2000US-0246474P.
PR
08-NOV-2000; 2000US-0246475P.
PR
08-NOV-2000; 2000US-0246476P.
PR
08-NOV-2000; 2000US-0246477P.
PR
08-NOV-2000; 2000US-0246478P.
PR
08-NOV-2000; 2000US-0246523P.
PR
08-NOV-2000; 2000US-0246524P.
PR
08-NOV-2000; 2000US-0246525P.
PR
08-NOV-2000; 2000US-0246526P.
PR
08-NOV-2000; 2000US-0246527P.
PR
08-NOV-2000; 2000US-0246528P.
PR
08-NOV-2000; 2000US-0246532P.
PR
08-NOV-2000; 2000US-0246609P.
PR
08-NOV-2000; 2000US-0246610P.
PR
08-NOV-2000; 2000US-0246611P.
PR
08-NOV-2000; 2000US-0246613P.
PR
17-NOV-2000; 2000US-0249207P.
PR
17-NOV-2000; 2000US-0249208P.
PR
17-NOV-2000; 2000US-0249209P.
PR
17-NOV-2000; 2000US-0249210P.
PR
17-NOV-2000; 2000US-0249211P.
PR
17-NOV-2000; 2000US-0249212P.
PR
17-NOV-2000; 2000US-0249213P.
PR
17-NOV-2000; 2000US-0249214P.
PR
17-NOV-2000; 2000US-0249215P.
PR
17-NOV-2000; 2000US-0249216P.
PR
17-NOV-2000; 2000US-0249217P.
PR
17-NOV-2000; 2000US-0249218P.
PR
17-NOV-2000; 2000US-0249244P.
PR
17-NOV-2000; 2000US-0249245P.
PR
17-NOV-2000; 2000US-0249264P.
PR
17-NOV-2000; 2000US-0249265P.
PR
17-NOV-2000; 2000US-0249297P.
PR
17-NOV-2000; 2000US-0249299P.
PR
17-NOV-2000; 2000US-0249300P.
PR
01-DEC-2000; 2000US-0250160P.
PR
05-DEC-2000; 2000US-0250391P.
PR
05-DEC-2000; 2000US-0251988P.
PR
05-DEC-2000; 2000US-0251989P.
PR
06-DEC-2000; 2000US-0251479P.
PR
08-DEC-2000; 2000US-0251856P.
PR
08-DEC-2000; 2000US-0251868P.
PR
08-DEC-2000; 2000US-0251869P.
PR
08-DEC-2000; 2000US-0251899P.
PR
11-DEC-2000; 2000US-0251990P.
PR
05-JAN-2001; 2001US-0259678P.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 40106; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 39068 BP; 10474 A; 9094 C; 8680 G; 10820 T; 0 U; 0 Other;
SQ

Query Match 25.8%; Score 86.4; DB 4; Length 39068;
Best Local Similarity 56.8%; Pred. No. 1.5e-07;
Matches 159; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 56 TGTITGGGAATTTAAACTACTGGTTTAAATTCCTTTTATTTCTTTTGTGTTGTAATGACTATT 115
Db 32919 TATATATACATATATACATATATATATATATATATATATATATATATATATATATAC 32860

QY 116 TTACGTCATATACAAATTTTATTTGTTTGTAAATGACTTTATTTGTTGTCATATGATAA 175
Db 32859 TATATACATATATACATATATATATATATATATATATATATATATATATATATACAT 32800

QY 176 TTTTATGTCATAGAACAAATTTTATTTGCTTGCATATATGACTTTATTTGTTATATGCTTATA 235
Db 32799 TACATACATATACAT 32740

QY 236 CAACAGATATTTTGTGTTTGTGACCGAGTCTTACTCTGTCTACCCAGGCTGGAGTGA 295
Db 32739 TATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 32680

QY 296 ATGGCAGTGTCTCAGCTCACTGCAACCTCCGCTCCCGGG 335
Db 32679 GTGGCAGACTCTCAGCTCACTGCACTGCTGCTGCTCCCTCCAGG 32640

RESULT 14
AAK73078
ID AAK73078 standard; DNA; 39068 BP.
XX
AC AAK73078;
XX
XX 06-NOV-2001 (first entry)
DT Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27890.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
PN
XX

PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225113P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
XX 26-SEP-2000; 2000US-0235484P.
XX 27-SEP-2000; 2000US-0235834P.
XX 27-SEP-2000; 2000US-0235836P.
XX 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.

29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX Disclosure; SEQ ID NO 27890; 3071pp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX SQ Sequence 39068 BP; 10820 A; 8680 C; 9094 G; 10474 T; 0 U; 0 Other;

Query Match 25.8%; Score 86.4; DB 4; Length 39068;
 Best Local Similarity 56.8%; Pred. No. 1.5e-07;
 Matches 159; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

Qy 56 TGGTTGTGGAAATTTAAACTACTGGTTTAATTTCTTTTATGTTGTTGAATATGACTATT 115
 Db 6150 TATATATACATATATACATATATATATACATATATATATATATATATATATACA 6209
 Qy 116 TTACGTATATACAAATTTTATTTGTTGTTAAAGACTTTATTTGTTGTCATATGATA 175
 Db 6210 TATATACATATATACATATATATATATATATATATATATATATATATATATATA 6269
 Qy 176 TTTTATGTCTAGAACAAATTTTATTTGCTTCATATATGACTTTTATTTGTTATATGCTATA 235
 Db 6270 TACATACATATACATATATATACATATATATATATATATATATATATATATATATA 6329
 Qy 236 CAACATAGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 295
 Db 6330 TATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6389
 Qy 296 ATGGCATGCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 335
 Db 6390 GTGGCAGCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 6429

RESULT 15
 AAK87544
 ID AAK87544 standard; DNA; 39068 BP.
 XX AAK87544;
 XX 07-NOV-2001 (first entry)
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:42356.
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX KW Cytostatic; gene therapy; vaccine; metastasis; ds.
 XX OS Homo sapiens.
 XX WO200157182-A2.
 XX 09-AUG-2001.
 XX 17-JAN-2001; 2001WO-US001354.
 XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226581P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0231145P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234397P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX XX
(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.
Disclosure; SEQ ID NO 42356; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX

SQ Sequence 39068 BP; 10820 A; 8680 C; 9094 G; 10474 T; 0 U; 0 Other;

Query Match 25.8%; Score 86.4; DB 4; Length 39068;
 Best Local Similarity 56.8%; Pred. No. 1 5e-07;
 Matches 159; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY	56	TGTTTGTGGAAATTTTAAACTACTGGTTTAAATTTTATTTGGTTGTAATGACTATT	115
Db	6150	TATATATACATATATATACATATATATATATATATATATATATATATATATATACA	6209
QY	116	TTACGTCATATAACAATTTTATTTGTTGTTTAAATGACTTTATTTGTTGTCATATGATAA	175
Db	6210	TATATACATATATACATA	6269
QY	176	TTTATATGTCATAGAACAAATTTTATTTGTTGTAATGACTTTTATTTGTTATGCTATA	235
Db	6270	TACATACATATACATA	6329
QY	236	CAACTAGATTTTGTGTTTGTGTTTGAACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGTA	295
Db	6330	TATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	6389
QY	296	ATGGCATGGTCTCAGCTCACTGCAACCTCCGCTCCGGG	335
Db	6390	GTGGCACAGTCTCAGCTCACTGCACTCTCTGCTCCCGAGG	6429

Search completed: May 21, 2004, 17:36:52
 Job time : 142.729 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 16:26:10 ; Search time 1344.31 Seconds
(without alignments)
10801.019 Million cell updates/sec

Title: US-09-914-651a-1

Perfect score: 335

Sequence: 1 aattatttttctttaacc.....tgcaactccgctcccggg 335

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.un.*

28: em.vi.*

29: em.htg.hum.*

30: em.htg.inv.*

31: em.htg.mus.*

32: em.htg.other.*

33: em.htg.pln.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	335	100.0	335	6	BD269371	BD269371 Controlli
2	335	100.0	93525	9	AF007544	AF007544 Homo sapi
3	335	100.0	117521	2	AC023784	AC023784 Homo sapi
4	335	100.0	129095	2	AC110742	AC110742 Homo sapi
5	335	100.0	158524	2	AL162372	AL162372 Homo sapi
C 6	335	100.0	187529	9	AC118273	AC118273 Homo sapi
7	327	97.6	1669	9	AF480875	AF480875 Homo sapi
8	327	97.6	135637	9	AP003400	AP003400 Homo sapi
9	327	97.6	137888	9	AP003122	AP003122 Homo sapi
C 10	327	97.6	192848	2	AC024234	AC024234 Homo sapi
11	222.8	66.5	246865	2	AC074003	AC074003 Homo sapi
12	100	29.9	161940	2	AC079289	AC079289 Homo sapi
13	99	28.6	125685	9	AC073310	AC073310 Homo sapi
C 14	95.8	28.6	184864	9	AC013553	AC013553 Homo sapi
15	95.6	28.5	163401	2	AC073251	AC073251 Homo sapi
16	95.6	28.5	167878	9	AC096645	AC096645 Homo sapi
17	95.2	28.4	91792	9	AC092355	AC092355 Homo sapi
C 18	94.8	28.3	161925	9	AC090103	AC090103 Homo sapi
19	94.6	28.2	103930	9	AL358234	AL358234 Human DNA
C 20	94.6	28.2	113075	9	AL512363	AL512363 Human DNA
21	94.6	28.2	190202	9	AC022021	AC022021 Homo sapi
22	94.6	28.2	192952	2	AC028966	AC028966 Homo sapi
C 23	94	28.1	113729	9	AF260011	AF260011 Homo sapi
24	94	28.1	166977	9	AC007198	AC007198 Homo sapi
C 25	94	28.1	340000	9	HS21C048	AL163248 Homo sapi
C 26	93.8	28.0	198935	2	AC068573	AC068573 Homo sapi
27	93.4	27.9	106954	9	AC104651	AC104651 Homo sapi
C 28	93.2	27.8	217985	9	AC021193	AC021193 Homo sapi
29	93	27.8	181413	2	AP001352	AP001352 Homo sapi
30	92.8	27.7	154639	2	AC090175	AC090175 Homo sapi
31	92.6	27.6	107415	9	AL513365	AL513365 Human DNA
C 32	92.4	27.6	169240	9	AC022069	AC022069 Homo sapi
33	92	27.5	103566	9	AC097481	AC097481 Homo sapi
34	91.6	27.3	87791	9	HSJ223D17	AL121982 Human DNA
35	91.6	27.3	167869	2	AC021339	AC021339 Homo sapi
36	91.6	27.3	213044	9	AC096576	AC096576 Homo sapi
C 37	91.6	27.3	240379	9	AC005077	AC005077 Homo sapi
38	91.4	27.3	143770	9	AL589987	AL589987 Human DNA
39	91.2	27.2	102910	9	AB041512	AB041512 Homo sapi
C 40	91.2	27.2	144234	2	AP002425	AP002425 Homo sapi
41	91.2	27.2	144279	2	AP001570	AP001570 Homo sapi
42	91.2	27.2	162217	9	AC068043	AC068043 Homo sapi
43	91.2	27.2	216972	9	AC067819	AC067819 Homo sapi
44	91	27.2	162500	9	AC104513	AC104513 Homo sapi
C 45	91	27.2	167434	9	AC079910	AC079910 Homo sapi

ALIGNMENTS

RESULT 1
BD269371
LOCUS
DEFINITION Controlling construct containing intron 3 of prostate-specific
membrane antigen gene.
ACCESSION BD269371
VERSION BD269371.1 GI:33079139
KEYWORDS JP 2002537807-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 335)
AUTHORS Molloy,P.L. and Watt,F.
TITLE Controlling construct containing intron 3 of prostate-specific

BD269371 335 bp DNA linear PAT 17-JUL-2003
Controlling construct containing intron 3 of prostate-specific

JOURNAL
membrane antigen gene
Patent: JP 2002537807-A 1 12-NOV-2002;
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION
COMMENT OS Homo sapiens (human)
PN JP 2002537807-A/1
PD 12-NOV-2002
PF 01-MAR-2000 JP 2000602768
PR 01-MAR-1999 AU PP 8956, 25-JAN-2000 AU PQ 5268 PI
PETER LAURENCE MOLLOY, FUSIKO WATT
PC C12N15/09, A61K38/00, A61K38/46, A61K48/00, A61P35/00, C12N5/10, PC
C12N15/00,
PC A61K37/02, A61K37/54, C12N5/00
CC Controlling construct containing intron
3 of prostate-specific
CC membrane
CC antigen gene
PH key Location/Qualifiers
FT source 1..335
/organism='Homo sapiens (human)'.
FEATURES
source
Location/Qualifiers
1..335
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 335; DB 6; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.1e-48;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATTATTTTTCCTTTAACTCTTCAACCTTCAAGAAACAGTGGCTTGACTCTGTTT 60
DB 1 AATTATTTTTCCTTTAACTCTTCAACCTTCAAGAAACAGTGGCTTGACTCTGTTT 60
QY 61 GTGAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 120
DB 61 GTGAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 120
QY 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTGTATATGATAATTTTA 180
DB 121 TCATATAACAATTTTATTGTTTGTAAATGACTTTATTGTTTGTATATGATAATTTTA 180
QY 181 TGTCATAGACAATTTTATTGTTTGTATATGACTTTATTGTTTGTATATGCTATCAACT 240
DB 181 TGTCATAGACAATTTTATTGTTTGTATATGACTTTATTGTTTGTATATGCTATCAACT 240
QY 241 AGATTTTGTGTTTGTGTTTGTGACGAGTCTTACTCTGTACCCAGGCTGAGTGTAAATGGC 300
DB 241 AGATTTTGTGTTTGTGTTTGTGACGAGTCTTACTCTGTACCCAGGCTGAGTGTAAATGGC 300
QY 301 ATGGTCTCAGTCACTGCAACCTCCGCTCCCGG 335
DB 301 ATGGTCTCAGTCACTGCAACCTCCGCTCCCGG 335

RESULT 2
AF007544
LOCUS
DEFINITION
Homo sapiens prostate-specific membrane antigen (PSM) gene,
complete cds.
AF007544 GI:2970122
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 93525)
O'Keefe,D.S., Su,S.L., Luo,Y., Horiguchi,Y., Baoich,D.J., Fair,W.R.,
Powell,C.T., Zandvliet,D., Russell,P.J., Molloy,P.L., Nowak,N.J.,
Shows,T.B., Mullins,C., Vonder Haar,R.A., Fair,W.R. and Heston,W.D.
Mapping, genomic organization and promoter analysis of the human
prostate-specific membrane antigen gene

Biochim. Biophys. Acta 1443 (1-2), 113-127 (1998)
99057588
9838072
2 (bases 1682 to 65315)
Heston,W.D.W., Su,S.L., Luo,Y., Huryk,R., Baoich,D.J., Fair,W.R.,
Mullins,C. and Vonder Haar,R.A.
Direct Submission
Submitted (06-JUN-1997) Memorial Sloan Kettering Cancer Center,
1275 York Avenue, New York, New York 10021, USA
2 (bases 1 to 93525)
O'Keefe,D.S., Su,S.L., Luo,Y., Horiguchi,Y., Baoich,D.J.,
Powell,C.T., Zandvliet,D., Russell,P.J., Molloy,P.L., Nowak,N.J.,
Mullins,C., Vonder Haar,R.A., Fair,W.R. and Heston,W.D.W.
Direct Submission
Submitted (18-MAR-1998) Memorial Sloan Kettering Cancer Center,
1275 York Avenue, New York, New York 10021, USA
Sequence update by submitter
On Mar 19, 1998 this sequence version replaced gi:2897945.
FEATURES
Location/Qualifiers
1..93525
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p11-12"
2488..64518
/genes="PSM"
join(2488..2863,4994..5099,10726..10912,18275..18376,
24400..24525,25314..25500,27927..28020,35216..35314,
36196..36281,37697..37816,39896..39978,41911..41974,
46402..46469,53129..53220,54364..54454,56661..56925,
57226..57307,62423..62515,64209..64518)
/genes="PSM"
/product="prostate-specific membrane antigen"
join(2745..2863,4994..5099,10726..10912,18275..18376,
24400..24525,25314..25500,27927..28020,35216..35314,
36196..36281,37697..37816,39896..39978,41911..41974,
46402..46469,53129..53220,54364..54454,56661..56925,
57226..57307,62423..62515,64209..64398)
/genes="PSM"
/note="PSMA"
/codon_start=1
/product="prostate-specific membrane antigen"
/protein_id="AAC83972.1"
/db_xref="GI:2897946"
/translation="MNLLHETDSAVATARRPRLWCAGALVLGGFFLLQFLRGWPK
SNEATNTPKNNKAFDLKAEINKKFLHNFTQIPHLAGTEQNFQAKIQOSQWKE
FGLDSVELAHYDVLISYNNKTHPNYISINEDGNEIFNTSLPEPPPGYENVSIDVPP
FAFSQGMPEGLDLYVNTARTEDFFKLRDMKNCISKIYARIYKVFNGKVKNAQ
LAGAGVILYSDPADYFAPGVKSYDPGNLPGGVQVQGNILNGLAGDPLTPGYANE
YARRGIAEAVGLSPVHPHIGYDAQLLEKMGSSAPPDSSWRGSLKVPYVNGPFT
GNFTOKVWHIHSNEVTRIYNVIGTLRGVDPDRYVILGHRDSWFGIDPQGA
AVVHEIVRSFGTLKKEGWRPRTILFASMDAREFGLLGSTEWAEENRLLQERGVAVI
NADSLIEGNYTLRVDCITPLMYSLVHNLTKELSPDEGFGKSLYESWTKSPSPFSG
MPRIKLGSGNDFEVFFQRLGASGRARYKKNWETNKFSGYDPLHSHVETIELEKPY
DPMFYHLTVAQVQGMVPELANSIVLPFCRDYAVLRYKDIYSIKMHPQEMKT
YVSFDSLFSVKNFETIASKFSEQLQEDKSNPIVLRMMNDQLMELEAFIDPLGIP
DRPFYRHVIYAPSSHNKVGESFPGIYDALFDIESKVPDSKAWGEVKRQIYVAFTVQ
AAATLSEVA"

ORIGIN

Query Match 100.0%; Score 335; DB 9; Length 93525;
Best Local Similarity 100.0%; Pred. No. 3.2e-49;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATTATTTTTCCTTTAACTCTTCAACCTTCAAGAAACAGTGGCTTGACTCTGTTT 60
DB 14760 AATTATTTTTCCTTTAACTCTTCAACCTTCAAGAAACAGTGGCTTGACTCTGTTT 14819
QY 61 GTGAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 120
DB 14820 GTGAAATTTTAACTACTGTTTAAATTTCTTTTATTGTTGTAATGACTATTTTACG 14879

```

Qy 121 TCATATACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 180
Db 14860 TCATATACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 14939
Qy 181 TGTATAGAACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 240
Db 14940 TGTATAGAACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 14999
Qy 241 AGATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 300
Db 15000 AGATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 15059
Qy 301 ATGCTCTCAGCTCACTGCAACTCCGCTCCGGG 335
Db 15060 ATGCTCTCAGCTCACTGCAACTCCGCTCCGGG 15094

RESULT 3
LOCUS AC023784 117521 bp DNA linear HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 11 clone CTD-2026G22, WORKING DRAFT
ACCESSION AC023784
VERSION AC023784.3 GI:7230809
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 117521)
Waterston,R.H.
The sequence of Homo sapiens clone
2 (bases 1 to 117521)
Waterston,R.H.
Direct Submission
Submitted (17-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Mar 13, 2000 this sequence version replaced gi:7021753.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H MS2026G22
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-Primer ET; 100% of reads
Chemistry: Dye-Terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 108841 bases at least Q40
Consensus quality: 111448 bases at least Q30
Consensus quality: 112731 bases at least Q20
Insert size: 132000; agarose-fp
Insert size: 115321; sum-of-ctnigs
Quality coverage: 3.18 in Q20 bases; agarose-fp
Quality coverage: 3.67 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1436: contig of 1436 bp in length
* 1437 1536: gap of unknown length
* 1537 2618: contig of 1082 bp in length
* 2619 2718: gap of unknown length

```

FEATURES

```

source
1..117521
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="CTD-2026G22"

```

ORIGIN

```

Query Match      100.0%; Score 335; DB 2; Length 117521;
Best Local Similarity 100.0%; Pred. No. 3e-49;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTTTTCCTTTAACTTTTCAAACTCAAGGAAACCAAGTTGGCTTGACTCTGTTT 60
Db 25261 AATTATTTTTCCTTTAACTTTTCAAACTCAAGGAAACCAAGTTGGCTTGACTCTGTTT 25320

Qy 61 GTGAAAATTTTAACTACTGTTTAACTTTTATTTGTTGTTGTAATGACTATTTTACG 120
Db 25321 GTGAAAATTTTAACTACTGTTTAACTTTTATTTGTTGTTGTAATGACTATTTTACG 25380

Qy 121 TCATATACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 180
Db 25381 TCATATACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 25440

Qy 181 TGTATAGAACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 240
Db 25441 TGTATAGAACAAATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 25500

Qy 241 AGATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 300
Db 25501 AGATTTTATTTGTTTAAATGACTTTATTTGTTGTCATATGATAATTTTA 25560

```

QY 301 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 335
 Db 25561 ATGGTCTCAGCTCACTGCAACCTCCGCTCCCGG 25595

RESULT 4
 AC110742 129095 bp DNA linear PRI 02-JUN-2002
 LOCUS Homo sapiens chromosome 11, clone CTD-2026G22, complete sequence.
 AC110742
 AC110742.4 GI:21313758
 HTG.
 Homo sapiens (human)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 129095)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Li, G., Maclean, C.,
 Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
 Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (15-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 129095)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Li, G., Maclean, C.,
 Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
 Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (15-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 129095)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Li, G., Maclean, C.,
 Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
 Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (27-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
 JOURNAL
 TITLE
 AUTHORS
 COMMENT
 4 (bases 1 to 129095)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Li, G., Maclean, C.,
 Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
 Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 2, 2002 this sequence version replaced gi:20334649.
 All repeats were identified using RepeatMasker:
 Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L21820
 Center clone name: 2026_G22

FEATURES
 source
 1. 129095
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="CTD-2026G22"
 /clone_lib="CITD1 Human BAC"
 2. 261
 /rpt_family="LTR16A"
 1334. 1562
 /rpt_family="MIR"
 complement(2716. 2929)
 /rpt_family="L2"
 3055. 3336
 /rpt_family="AluJo"
 3468. 3524
 /rpt_family=" (TTCC)n"
 3525. 3560
 /rpt_family=" (TTCC)n"
 4640. 4849
 /rpt_family="MER20"
 complement(4884. 5129)
 /rpt_family="L3"
 complement(5208. 5413)
 /rpt_family="MIR"
 5840. 5891
 /rpt_family="L1MC5"
 complement(6065. 8047)
 /rpt_family="L1PA6"
 8050. 10029
 /rpt_family="L1PA6"
 10340. 10497
 /rpt_family="L1PA6"

```

/rept_family="L1ME4A"
10670..10691
/rept_family="AT_rich"
10977..10997
/rept_family="AT_rich"
complement(10998..11293)
/rept_family="AluJo"
12004..12042
/rept_family="AT_rich"
12337..12396
/rept_family="(TA)n"
12447..12489
/rept_family="AT_rich"
12950..12970
/rept_family="AT_rich"
complement(13047..13097)
/rept_family="MIR3"
13218..13339
/rept_family="MIR"
13460..13500
/rept_family="AT_rich"
14066..14087
/rept_family="AT_rich"
14141..14166
/rept_family="(TTTTTG)n"
15042..15277
/rept_family="AluSp"
complement(16069..16260)
/rept_family="MIR"
complement(16362..16449)
/rept_family="MIR"
16603..16637
/rept_family="(TG)n"
16976..17065
/rept_family="L1"
17264..17286
/rept_family="(A)n"
17817..17860
/rept_family="(TTTG)n"
18317..18432
/rept_family="L2"
18922..19089
/rept_family="MER45A"
20139..20250
/rept_family="FLAM_C"
20252..20292
/rept_family="(TAAAA)n"
20804..21817
/rept_family="L1MA3"
21815..22853
/rept_family="L1MA3"
22833..24145
/rept_family="L1MA3"
24826..24896
/rept_family="AT_rich"
25215..25380
/rept_family="MIR"
25381..25402
/rept_family="(TG)n"
25403..25440
/rept_family="MIR"
complement(26317..26407)

Query Match      100.0%; Score 335; DB 9; Length 129095;
Best Local Similarity 100.0%; Pred. No. 2.9e-49;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATATTTTCTTTAACTTTTCAAACTCAAGGAAACAGTTGGCTTGACTCTGTTT 60
Db 123370 AATATTTTCTTTAACTTTTCAAACTCAAGGAAACAGTTGGCTTGACTCTGTTT 123429
QY 61 GTGGAATTTTAACTGTTTAACTTTTCTTTATTTGTTGTAATGACTATTTTACG 120
|||||

```

Db 123430 GTGGAATTTTAACTGTTTAACTTTCTTTATTTGTTGTAATGACTATTTTACG 123489

QY 121 TCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTTCATATGATAATTTTA 180
|||||

Db 123490 TCATATAACAATTTTATTTGTTTAAATGACTTTATTTGTTTCATATGATAATTTTA 123549
|||||

QY 181 TGTATAGAACAAATTTTATTTGTTTATGATATATGACTTTATTTGTTTATGCTATACAACT 240
|||||

Db 123550 TGTATAGAACAAATTTTATTTGTTTATGATATATGACTTTATTTGTTTATGCTATACAACT 123609
|||||

QY 241 AGATTTTGTGTTTGTGTTTACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGAATGCG 300
|||||

Db 123610 AGATTTTGTGTTTGTGTTTACCGAGTCTTACTCTGTCACCCAGGCTGGAGTGAATGCG 123669
|||||

QY 301 ATGCTCTCAGCTCACTGCAACCTCGCTCCGGG 335
|||||

Db 123670 ATGCTCTCAGCTCACTGCAACCTCGCTCCGGG 123704
|||||

RESULT 5

AL162372

LOCUS Homo sapiens chromosome 13 clone RP11-181P16, 3 unordered pieces. 158524 bp DNA linear HTG 10-JUL-2001

DEFINITION AL162372

ACCESSION AL162372.11 GI:12539555

VERSION HTG; HTGS PHASE1; HTGS_CANCELLED.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Burton, J.

REFERENCE 1

AUTHORS Direct Submission

TITLE Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

JOURNAL requests: clonerequest@sanger.ac.uk

On Jan 25, 2001 this sequence version replaced gi:11693478.

COMMENT ----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: hA181P16

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: pGAP4; 108752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 157639 bases at least Q40

Consensus quality: 157977 bases at least Q30

Consensus quality: 158153 bases at least Q20

Insert size: 158324; sum-of-contigs

Insert size: 174362; 5.7% error; agarose-fp

Quality coverage: 6.90x in Q20 bases; sum-of-contigs Quality coverage: 6.26x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 59356: contig of 59356 bp in length

* 59357 59456: gap of 100 bp

* 59457 65248: contig of 5792 bp in length

* 65249 65348: gap of 100 bp

* 65349 158524: contig of 93176 bp in length.

Location/Qualifiers

1..158524

/organism="Homo sapiens"

/mol_type="genomic DNA"

FEATURES

source

```
/db_xref="taxon:9606"
/chromosomes="13"
/clone="RP11-181P16"
/clone_lib="RP11-11.1"
I. .59356
/note="assembly_fragment:00306
clone end:T7
vector side:left"
59457_-65248
/note="assembly_fragment:01282
fragment_chain:1"
65349_-158524
/note="assembly_fragment:02815
fragment_chain:1"

ORIGIN
Query Match 100.0%; Score 335; DB 2; Length 158524;
Best Local Similarity 100.0%; Pred. No. 2.8e-49;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAAACAGTGGCTTGACTCTGTTT 60
Db 71493 AATTATTTTTCCTTTAACTTTCAAACTCAAGGAAAACAGTGGCTTGACTCTGTTT 71552

QY 61 GTGAAAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTTAATGACTATTTTACG 120
Db 71553 GTGAAAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTTAATGACTATTTTACG 71612

QY 121 TCATATAACAATTTTATTTGTTTAAAGACTTTATTTGTTTGTCTATGATAATTTTA 180
Db 71613 TCATATAACAATTTTATTTGTTTAAAGACTTTATTTGTTTGTCTATGATAATTTTA 71672

QY 181 TGCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTATATGCTATACAACT 240
Db 71673 TGCATAGAACAAATTTTATTTGTTGATATATGACTTTATTTGTTATATGCTATACACT 71732

QY 241 AGATTTTGTGTTTGTGTTTACCGAGCTTACTGTGTCACCGAGCTGGAGTGTATGCG 300
Db 71733 AGATTTTGTGTTTGTGTTTACCGAGCTTACTGTGTCACCGAGCTGGAGTGTATGCG 71792

QY 301 ATGCTCTCAGTCACTGCACTCCGCTCCGCTCCGCG 335
Db 71793 ATGCTCTCAGTCACTGCACTCCGCTCCGCTCCGCG 71827

RESULT 6
AC118273/c 187529 bp DNA linear PRI 09-SEP-2002
LOCUS Homo sapiens chromosome 11, clone RP11-107P7, complete sequence.
DEFINITION AC118273
ACCESSION AC118273
VERSION AC118273.4 GI:22759265
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 187529)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Boukhgalter,B., Brown,A., Camarata,J., Chang,J., Chazaro,B., Collymore,A.,
Chang,J., Chazaro,B., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Gardyna,S., Gordon,S., Horton,L., Hulme,W., Johnson,R., Jones,C., Kamat,A.,
Karatlas,A., Kells,C., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Roman,J.,
Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Sep 9, 2002 this sequence version replaced gi:22539138.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26235
Center clone name: 107_P_7
-----
Location/Qualifiers
1. 187529
/organism="Homo sapiens"
FEATURES
SOURCE
```

```
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
O'Neil,D., Peterson,K., Roman,J., Roy,A., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT 3 (bases 1 to 187529)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Boukhgalter,B., Brown,A., Camarata,J., Chang,J., Chazaro,B., Collymore,A.,
Chang,J., Chazaro,B., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Gardyna,S., Gordon,S., Horton,L., Hulme,W., Johnson,R., Jones,C., Kamat,A.,
Karatlas,A., Kells,C., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Roman,J.,
Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT 4 (bases 1 to 187529)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Boukhgalter,B., Brown,A., Camarata,J., Chang,J., Chazaro,B., Collymore,A.,
Chang,J., Chazaro,B., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Gardyna,S., Gordon,S., Horton,L., Hulme,W., Johnson,R., Jones,C., Kamat,A.,
Karatlas,A., Kells,C., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Roman,J.,
Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Sep 9, 2002 this sequence version replaced gi:22539138.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26235
Center clone name: 107_P_7
-----
Location/Qualifiers
1. 187529
/organism="Homo sapiens"
FEATURES
SOURCE
```


I-77-22/100, 1000000-1000000, 1000000-1000000, 1000000-1000000, 1000000-1000000

THESE ANSWERS ARE FROM /


```
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-31312"

Query Match      97.6%; Score 327; DB 9; Length 137888;
Best Local Similarity 98.5%; Pred. No. 7e-48;
Matches 330; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

ORIGIN
1 AATTATTTTTCCTTTAACTTCAAACTCAAGGAAACCAAGTTCGCTTCACTCTGTTT 60
825 AATTATTTTTCCTTTAACTTCAAACTCAAGGAAACCAAGTTCGCTTCACTCTGTTT 884
61 GTGGAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTAATATGACTATTTTACG 120
885 GTGGAATTTTAACTACTGTTTAAATTTCTTTATTTGTTGTAATATGACTATTTTACG 944
121 TCATATACAAATTTTATTTGTTTAAATGACATTTATTTGTTGTCATATGATATTTTA 180
945 TCATATACAAATTTTATTTGTTTAAATGACATTTATTTGTTGTCATATGATATTTTA 1004
181 TGTATAGAACAAATTTTATTTGTTTAAATGACATTTATTTGTTGTAATATGACTATTTTAC 240
1005 TGTATAGAACAAATTTTATTTGTTTAAATGACATTTATTTGTTGTAATATGACTATTTTAC 1064
241 AGATTTTGTGTTTGTGTTTACCGAGTCTTACTCTGTCCACCCAGGCTGGAGTGAATGGC 300
1065 AGATTTTGTGTTTGTGTTTACCGAGTCTTACTCTGTCCACCCAGGCTGGAGTGAATGGC 1124
301 ATGCTCTCAGCTCACTGCAACTCGGCTCCCGGG 335
1125 ATGCTCTCAGCTCACTGCAACTCGGCTCCCGAG 1159

RESULT 10
AC024234/5 192648 bp DNA linear HTG 01-SEP-2000
LOCUS Homo sapiens chromosome 11 clone RP11-31312, WORKING DRAFT
DEFINITION AC024234
ACCESSION AC024234.5 GI:8569915
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 192648)
Waterston, R.H.
Direct Submission
Submitted (28-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 16, 2000 this sequence version replaced gi:7284696.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0311302
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing method: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 181493 bases at least Q40
Consensus quality: 184864 bases at least Q30

----- Location/Qualifiers -----
1..192648
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"

FEATURES
source
```

Consensus quality: 186864 bases at least Q20
Insert size: 18000; agarose-fp
Insert size: 190048; sum-of-contigs
Quality coverage: 3.74 in Q20 bases; agarose-fp
Quality coverage: 3.73 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1413: contig of 1413 bp in length
1513: gap of unknown length
1514 1978: contig of 1465 bp in length
2979 3078: gap of unknown length
3079 4859: contig of 1781 bp in length
4860 4959: gap of unknown length
4960 6723: contig of 1764 bp in length
6724 9068: gap of unknown length
9069 9169: contig of 2245 bp in length
9170 11510: contig of 2342 bp in length
11511 11610: gap of unknown length
11611 15431: contig of 3821 bp in length
15432 15532: gap of unknown length
15533 18469: contig of 2938 bp in length
18470 22092: contig of 3523 bp in length
22093 22192: gap of unknown length
22193 25250: contig of 3058 bp in length
25251 28352: contig of 3002 bp in length
28353 28452: gap of unknown length
28453 31550: contig of 3097 bp in length
31551 34631: contig of 2982 bp in length
34632 34732: gap of unknown length
34733 38568: contig of 3737 bp in length
38569 42337: contig of 3769 bp in length
42338 42438: gap of unknown length
42439 47543: contig of 5106 bp in length
47544 56048: gap of unknown length
56049 64671: contig of 8405 bp in length
64672 73100: contig of 8523 bp in length
73101 73200: gap of unknown length
73201 81814: contig of 8329 bp in length
81815 81914: gap of unknown length
81915 91133: contig of 8614 bp in length
91134 91234: gap of unknown length
91235 101427: contig of 9219 bp in length
101428 101527: contig of 10194 bp in length
101528 13041: gap of unknown length
13042 13141: contig of 11514 bp in length
13142 125504: gap of unknown length
125505 125604: contig of 12363 bp in length
125605 140510: gap of unknown length
140511 140610: contig of 14906 bp in length
140611 158204: gap of unknown length
158205 158305: contig of 17594 bp in length
158306 192648: gap of unknown length
192649 34344: contig of 34344 bp in length.

```
/clone="RP11-31312"
1. .1413
/note="assembly_name:Contig10"
1514. .2978
/note="assembly_name:Contig11"
3079. .4859
/note="assembly_name:Contig12"
4960. .6723
/note="assembly_name:Contig13"
6824. .9068
/note="assembly_name:Contig14"
9169. .11510
/note="assembly_name:Contig15"
11611. .15431
/note="assembly_name:Contig16"
15532. .18469
/note="assembly_name:Contig17"
18570. .22092
/note="assembly_name:Contig18"
22193. .25250
/note="assembly_name:Contig19"
25351. .28352
/note="assembly_name:Contig20"
28453. .31549
/note="assembly_name:Contig21"
31650. .34631
/note="assembly_name:Contig22"
34732. .38468
/note="assembly_name:Contig23"
38569. .42337
/note="assembly_name:Contig24"
42438. .47543
/note="assembly_name:Contig25"
47644. .56048
/note="assembly_name:Contig26"
56149. .84671
/note="assembly_name:Contig27"
64772. .73100
/note="assembly_name:Contig28"
73201. .81814
/note="assembly_name:Contig29"
81915. .91133
/note="assembly_name:Contig30"
91234. .101427
/note="assembly_name:Contig31"
clone end:SP6
vector side:left"
101528. .113041
/note="assembly_name:Contig32"
113142. .125504
/note="assembly_name:Contig33"
125605. .140510
/note="assembly_name:Contig34"
140611. .158204
/note="assembly_name:Contig35"
158305. .192648
/note="assembly_name:Contig36"

ORIGIN
Query Match 97.6%; Score 327; DB 2; Length 192648;
Best Local Similarity 98.5%; Pred. No. 6.5e-48;
Matches 330; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AATTATTTTTCCTTTAACTCTCAAACTCAAGGAAACACGATGGCTTGACTCTCTTT 60
DB 174244 AATTATTTTTCCTTTAACTCTCAAACTCAAGGAAACACGATGGCTTGACTCTCTTT 174185

QY 61 GTGAAATTTTAACTACTGGTTTAACTTTTATTTGTTGTAATGACTATTTTACG 120
DB 174184 GTGAAATTTTAACTACTGGTTTAACTTTTATTTGTTGTTATATGACTATTTTATG 174125

QY 121 TCATATACATTTTATTTTCTTTGTTTAAATGACTTTATTTGTTCTCATATGATAATTTTA 180
```

```
* 35133 50532: contig of 15400 bp in length
* 50533 50632: gap of unknown length
* 50633 64756: contig of 14124 bp in length
* 64757 64856: gap of unknown length
* 64857 80978: contig of 16122 bp in length
* 80979 81078: gap of unknown length
* 81079 101159: contig of 20081 bp in length
* 101160 10259: gap of unknown length
* 10259 102865: contig of 1606 bp in length
* 102866 102965: gap of unknown length
* 102966 104593: contig of 1628 bp in length
* 104594 104693: gap of unknown length
* 104694 106231: contig of 1338 bp in length
* 106232 106331: gap of unknown length
* 106332 107693: contig of 1362 bp in length
* 107694 107793: gap of unknown length
* 107794 109155: contig of 1362 bp in length
* 109156 109255: gap of unknown length
* 109256 110481: contig of 1226 bp in length
* 110482 110581: gap of unknown length
* 110582 111881: contig of 1300 bp in length
* 111882 111981: gap of unknown length
* 111982 113524: contig of 1543 bp in length
* 113525 113624: gap of unknown length
* 113625 114983: contig of 1359 bp in length
* 114984 115083: gap of unknown length
* 115084 116145: contig of 1062 bp in length
* 116146 116245: gap of unknown length
* 116246 117625: contig of 1380 bp in length
* 117626 117725: gap of unknown length
* 117726 120009: contig of 2284 bp in length
* 120010 120109: gap of unknown length
* 120110 122469: contig of 2360 bp in length
* 122470 122569: gap of unknown length
* 122570 124001: contig of 1432 bp in length
* 124002 124101: gap of unknown length
* 124102 126352: contig of 2251 bp in length
* 126353 126452: gap of unknown length
* 126453 128477: contig of 2025 bp in length
* 128478 128577: gap of unknown length
* 128578 130466: contig of 1889 bp in length
* 130467 130566: gap of unknown length
* 130567 132508: contig of 1939 bp in length
* 132509 132605: gap of unknown length
* 132606 134148: contig of 1543 bp in length
* 134149 134248: gap of unknown length
* 134249 136448: contig of 2200 bp in length
* 136449 136548: gap of unknown length
* 136549 138255: contig of 1707 bp in length
* 138256 138355: gap of unknown length
* 138356 140535: contig of 2180 bp in length
* 140536 140635: gap of unknown length
* 140636 142199: contig of 1564 bp in length
* 142200 142299: gap of unknown length
* 142300 144900: contig of 2607 bp in length
* 144901 145008: gap of unknown length
* 145009 147213: contig of 2207 bp in length
* 147214 147313: gap of unknown length
* 147314 148878: contig of 1565 bp in length
* 148879 148978: gap of unknown length
* 148979 150908: contig of 1928 bp in length
* 150910 151008: gap of unknown length
* 151009 153354: contig of 2348 bp in length
* 153355 153454: gap of unknown length
* 153455 155071: contig of 1617 bp in length
* 155072 155171: gap of unknown length
* 155172 159010: contig of 3839 bp in length
* 159011 159110: gap of unknown length
* 159111 162070: contig of 2960 bp in length
* 162071 162170: gap of unknown length
* 162171 164452: contig of 2282 bp in length
* 164453 164552: gap of unknown length
* 164553 166294: contig of 3742 bp in length

* 168295 168394: gap of unknown length
* 168395 171385: contig of 2991 bp in length
* 171386 171485: gap of unknown length
* 171486 175016: contig of 3331 bp in length
* 175017 175116: gap of unknown length
* 175117 178766: contig of 3650 bp in length
* 178767 178866: gap of unknown length
* 178867 182838: contig of 3972 bp in length
* 182839 182938: gap of unknown length
* 182939 186148: contig of 3210 bp in length
* 186149 186248: gap of unknown length
* 186249 190168: contig of 3920 bp in length
* 190169 190268: gap of unknown length
* 190269 193524: contig of 3256 bp in length
* 193525 193624: gap of unknown length
* 193625 202160: contig of 8536 bp in length
* 202161 202260: gap of unknown length
* 202261 210133: contig of 7873 bp in length
* 210134 210233: gap of unknown length
* 210234 218141: contig of 7908 bp in length
* 218142 218241: gap of unknown length
* 218242 228213: contig of 9972 bp in length
* 228214 228313: gap of unknown length
* 228314 237496: contig of 9183 bp in length
* 237497 237597: gap of unknown length
* 237597 246865: contig of 9269 bp in length.

FEATURES
      Location/Qualifiers
      source
      1..246865
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="2"
      /clone="RP11-163G14"
      1..12900
      /note="assembly_name:Contig100"
      13001..23006
      /note="assembly_name:Contig101"
      23107..35032
      /note="assembly_name:Contig102"
      35133..50532
      /note="assembly_name:Contig103"
      50633..64756
      /note="assembly_name:Contig104"
      64857..80978
      /note="assembly_name:Contig105"
      81079..101159
      /note="assembly_name:Contig106"
      101260..102865
      /note="assembly_name:Contig107"
      102966..104593
      /note="assembly_name:Contig108"
      104694..106231
      /note="assembly_name:Contig109"
      106332..107693
      /note="assembly_name:Contig110"
      107794..109155
      /note="assembly_name:Contig111"
      109256..110481
      /note="assembly_name:Contig112"
      110582..111881
      /note="assembly_name:Contig113"
      111982..113524
      /note="assembly_name:Contig114"
      113625..114983
      /note="assembly_name:Contig115"
      115084..116145
      /note="assembly_name:Contig116"
      116246..117625
      /note="assembly_name:Contig117"
      117726..120009
      /note="assembly_name:Contig118"

Query Match      66.5%; Score 222.8; DB 2; Length 246865;
```



```
misc_feature 60141..61238
              /note="assembly_fragment"
misc_feature 61339..65067
              /note="assembly_fragment"
misc_feature 65168..107152
              /note="assembly_fragment"
misc_feature 107253..154782
              /note="assembly_fragment"
misc_feature 154883..161940
              /note="assembly_fragment"
              clone_end:T7
              vector_side:right"

ORIGIN
Query Match      29.9%; Score 100; DB 2; Length 161940;
Best Local Similarity 62.11%; Pred.No.1.2e-08;
Matches 175; Conservative 0; Mismatches 105; Indels 2; Gaps 1;

QY 54 TCTGTTTGGGAAATTTTAAACTACTGCTTTTAAATTTCTTTATTTGGTTGCTATATGACTTA 113
Db 91637 TTTATATATTTATATATGTTTATATATATTTATATATGTTTATATATGTTTATATATTTTA 91696

QY 114 TTTTACGTCATATACAAATTTTATTTGTTTAAAGACATTTTATTTGTTTGCATATGAT 173
Db 91697 TATATGTTTATATATATATTTATATATATGTTTATATATTTATATATATTTTATATTTT 91756

QY 174 AATTTTATGTCATAGAACAAATTTTATTTGTTTGTATATATGACTTTTATTTATATGCGCTA 233
Db 91757 TATATTTTATGTTTATATATATTTATATATTTTATATATTTTATATATTTTATATTTATA--TA 91814

QY 234 TACAACTAGATTTTCTGTTTGTGTTTACCGAGTCTACTCTGTCACCCAGCTGGAGTG 293
Db 91815 TATATATATATTTTATTTTATTTTGTGAGCCAGAGTCTGCTCTGTCACCCAGCTGGAGTG 91874

QY 294 TAATGGCATGCTCAGCTCACTGCAACTCCGCTCCCGG 335
Db 91875 CAGTGGCAGCATCTGGCTCACTACAGCTCCACTCCCGG 91916

RESULT 13
AC073310 LOCUS Homo sapiens BAC clone RP11-49G5 from 7, complete sequence.
DEFINITION Homo sapiens BAC clone RP11-49G5 from 7, complete sequence.
ACCESSION AC073310
VERSION AC073310.7 GI:13176604
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 125685)
AUTHORS Sulston,J.E. and Wilson,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 125685)
AUTHORS Haakenson,W., Hannah,C. and Kang,K.
TITLE The sequence of Homo sapiens BAC clone RP11-49G5
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 125685)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 125685)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 125685)
```

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Waterston,R.
Direct Submission
Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 125685)
Waterston,R.
Direct Submission
Submitted (10-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 125685)
Waterston,R.
Direct Submission
Submitted (26-APR-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
8 (bases 1 to 125685)
Wilson,R.
Direct Submission
Submitted (08-OCT-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 1, 2001 this sequence version replaced gi:11024926.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0049G05

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Moon,P.Y., Zhao,B., Frengen,E., Tatenio,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-409M7, 200 bp overlap the clone sequenced to the right is RP11-374N8. Actual start of this clone is at base position 14572 of RP11-409M7 actual end is at base position 125685 of RP11-49G5.

FEATURES

Location/Qualifiers
1..125685
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"

Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 184864)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Menes, L., Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-AUG-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Dec 13, 2000 this sequence version replaced gi:9280765.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2203

Center clone name: 325_L_12

FEATURES

source

Location/Qualifiers

1. .184864
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RPL1-325L12"
/clone_lib="RPL1-11 Human Male BAC"

286. .542

/rpt_family="AluJo"
complement(1302. .1593)

/rpt_family="AluSc"

1769. .2080

/rpt_family="AluSx"

2099. .2161

/rpt_family="L2"

2215. .2322

/rpt_family="L2"

2673. .2700

/rpt_family="AT_rich"

2742. .2856

/rpt_family="MIR3"

2888. .3180

/rpt_family="AluSp"

3289. .3597

/rpt_family="AluY"
3936. .4107
/rpt_family="FAM"
4158. .4204
/rpt_family="L2"
4379. .4492
/rpt_family="MIR"
complement(4500. .4543)
/rpt_family="MIR"
complement(4544. .4846)
/rpt_family="AluSx"
complement(4847. .4972)
/rpt_family="MIR"
5225. .5333
/rpt_family="MIR3"
5363. .5655
/rpt_family="L1MASA"
complement(5656. .5781)
/rpt_family="MLT2D"
5790. .6207
/rpt_family="L1MASA"
complement(6313. .6695)
/rpt_family="MLT2B1"
complement(6696. .6998)
/rpt_family="AluJb"
complement(6999. .7108)
/rpt_family="MLT2B1"
complement(7117. .7273)
/rpt_family="MIR"
complement(7485. .7608)
/rpt_family="MER5A"
complement(8299. .8609)
/rpt_family="AluSp"
complement(9109. .9192)
/rpt_family="MIR"
complement(9302. .9643)
/rpt_family="L1PA10"
9647. .9916
/rpt_family="L2"
10055. .10271
/rpt_family="L2"
complement(10361. .10480)
/rpt_family="MER94"
10715. .11025
/rpt_family="AluJo"
11252. .11563
/rpt_family="AluY"
11818. .12119
/rpt_family="AluSg"
complement(12564. .12623)
/rpt_family="WADE1"
12710. .13002
/rpt_family="AluSg"
12967. .12977
/note="<30 qual SNGL region"
13004. .13008
/note="<30 qual SNGL region"
complement(13154. .13452)
/rpt_family="AluSp"
complement(14745. .15042)
/rpt_family="AluSx"
complement(15514. .15810)
/rpt_family="AluY"
16087. .16465
/rpt_family="MER61D"
complement(16517. .16762)
/rpt_family="L1MC4a"
complement(17245. .17342)
/rpt_family="L1MB1"
complement(17599. .17737)
/rpt_family="MIR3"
18456. .18775
/rpt_family="L2"

unsure

unsure

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

Mon May 24 08:12:32 2004

```
repeat_region complement(18776..19087)
repeat_region /rpt_family="AluJb"
repeat_region 19088..19313
repeat_region /rpt_family="L2"
repeat_region complement(19342..20175)
repeat_region /rpt_family="L1MB7"
repeat_region 20176..20477
repeat_region /rpt_family="AluSx"
repeat_region complement(20478..20620)
repeat_region /rpt_family="L1MB7"
repeat_region complement(20621..20931)
repeat_region /rpt_family="AluJb"
repeat_region complement(20932..21437)
repeat_region /rpt_family="L1MB7"
repeat_region 21438..21459
repeat_region /rpt_family="(TAA)n"
repeat_region complement(21460..21545)
repeat_region /rpt_family="L1MB7"
repeat_region complement(21546..21834)
repeat_region /rpt_family="AluSx"
repeat_region complement(21835..22416)
repeat_region /rpt_family="L1MB7"
repeat_region complement(22417..22548)
repeat_region /rpt_family="FLAM A"
repeat_region complement(22549..22796)
repeat_region /rpt_family="L1MB7"
repeat_region 22797..23107
repeat_region /rpt_family="AluSx"
repeat_region complement(23108..23247)
repeat_region /rpt_family="L1MB7"
repeat_region 23255..23457

Query Match 28.6%; Score 95.8; DB 9; Length 184864;
Best Local Similarity 57.3%; Pred. No. 6e-08;
Matches 192; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 2 ATTATTTTTCCTTAACTCTGTTTAACTTCAAACTCAAGGAAACCCAGTGGCGCTTGACTCTGTTG 61
Db 155318 ATTATCTTAAATTTATTAATTTTAAATGATCCATATAATGCAATATTTTATTTTATATA 155259

QY 62 TGGAAATTTTAACTACTGTTTAACTTCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 121
Db 155258 TATATATTTTATATATATATTTTATATATATATATATATATATATATATATATATATATAT 155199

QY 122 CATATAACAATTTTATTTGTTTAAATGACTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 181
Db 155198 TATATATTTTATATATATATTTTATATATATATATATATATATATATATATATATATATATA 155139

QY 182 GTCATAGAACAAATTTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240
Db 155138 TATATATTTTATATATATATTTTATATATATATATATATATATATATATATATATATATAT 155079

QY 241 AGATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
Db 155078 ATATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 155019

QY 301 ATGGTCTCAGTCTCACTCAACTCCGCTCCCGGG 335
Db 155018 CGGATCTCAGTCTCACTCAAGCTCCACCTCCAGG 154984
```

```
RESULT 15
AC073251
LOCUS AC073251 163401 bp DNA linear HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-102D20, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
ACCESSION AC073251
VERSION AC073251.2 GI:8954214
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 163401)
Waterson,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 163401)
Waterson,R.H.
Direct Submission
Submitted (12-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 7, 2000 this sequence version replaced gi:8468953.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0102D20
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 145004 bases at least Q40
Consensus quality: 150190 bases at least Q30
Consensus quality: 153173 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 160201; sum-of-contigs
Quality coverage: 3.36 in Q20 bases; agarose-fp
Quality coverage: 3.59 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1128: contig of 1128 bp in length
1128: gap of unknown length
1229: contig of 1293 bp in length
2522: gap of unknown length
2622: contig of 1359 bp in length
3981: gap of unknown length
4081: contig of 2197 bp in length
6277: gap of unknown length
6377: contig of 2122 bp in length
8499: gap of unknown length
8599: contig of 2287 bp in length
10886: gap of unknown length
10887: contig of 2748 bp in length
13734: contig of 2504 bp in length
13834: gap of unknown length
16974: contig of 3140 bp in length
17074: gap of unknown length
17075: contig of 3713 bp in length
20887: gap of unknown length
20888: contig of 2504 bp in length
23391: contig of 3757 bp in length
23492: gap of unknown length
27249: contig of 3400 bp in length
27349: gap of unknown length
30849: contig of 3280 bp in length
34129: gap of unknown length
34229: contig of 3047 bp in length
37275: gap of unknown length
37276: contig of 3333 bp in length
37376: gap of unknown length
40709: contig of 3236 bp in length
40809: gap of unknown length
44045: contig of 3795 bp in length
47939: contig of 3795 bp in length


```

* 47940 48039: gap of unknown length
* 48040 53255: contig of 5216 bp in length
* 53256 53355: gap of unknown length
* 53356 57715: contig of 4360 bp in length
* 57716 57815: gap of unknown length
* 57816 61875: contig of 4060 bp in length
* 61876 61975: gap of unknown length
* 61976 67647: contig of 5672 bp in length
* 67648 67748: gap of unknown length
* 67749 73739: contig of 5892 bp in length
* 73740 78458: contig of 4719 bp in length
* 78459 84299: gap of unknown length
* 84300 84399: contig of 5741 bp in length
* 84400 88820: gap of unknown length
* 88821 88920: contig of 4421 bp in length
* 88921 96269: gap of unknown length
* 96270 96369: contig of 7349 bp in length
* 96370 103347: contig of 6978 bp in length
* 103348 103447: gap of unknown length
* 103448 109631: contig of 6184 bp in length
* 109632 109731: gap of unknown length
* 109732 118014: contig of 8283 bp in length
* 118015 118114: gap of unknown length
* 118115 129041: contig of 10927 bp in length
* 129042 129141: gap of unknown length
* 129142 139215: contig of 10074 bp in length
* 139216 139315: gap of unknown length
* 139316 151356: contig of 12041 bp in length
* 151357 151457: gap of unknown length
* 151457 163401: contig of 11945 bp in length.
FEATURES
    source
        1..163401
            /organism="Homo sapiens"
            /mol_type="Genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clone="RP11-102D20"
        1..11128
            /note="assembly_name:Contig9"
            1229..2521
                /note="assembly_name:Contig10"
            2622..3980
                /note="assembly_name:Contig11"
            4081..6277
                /note="assembly_name:Contig12"
            6378..8499
                /note="assembly_name:Contig13"
            8600..10886
                /note="assembly_name:Contig14"
            10987..113734
                /note="assembly_name:Contig15"
            13835..16974
                /note="assembly_name:Contig16"
            17075..20787
                /note="assembly_name:Contig17"
            20888..23391
                /note="assembly_name:Contig18"
            clone end:SP6
            vector_side:left
            23492..27248
                /note="assembly_name:Contig19"
            27349..30748
                /note="assembly_name:Contig20"
            30849..34128
                /note="assembly_name:Contig21"
            34229..37275
                /note="assembly_name:Contig22"
            37376..40708
                /note="assembly_name:Contig23"
            40809..44044
                /note="assembly_name:Contig24"

```

```

misc_feature 44145..47939
              /note="assembly_name:Contig25"
misc_feature 48040..53255
              /note="assembly_name:Contig26"
misc_feature 53356..57715
              /note="assembly_name:Contig27"
misc_feature 57816..61875
              /note="assembly_name:Contig28"
misc_feature 61976..67647
              /note="assembly_name:Contig29"
misc_feature 67748..73639
              /note="assembly_name:Contig30"
misc_feature 73740..78458
              /note="assembly_name:Contig31"
misc_feature 78559..84299
              /note="assembly_name:Contig32"
misc_feature 84400..88820
              /note="assembly_name:Contig33"
misc_feature 88921..96269
              /note="assembly_name:Contig34"
misc_feature 96370..103347
              /note="assembly_name:Contig35"
misc_feature 103448..109631
              /note="assembly_name:Contig36"
misc_feature 109732..118014
              /note="assembly_name:Contig37"
misc_feature 118115..129041
              /note="assembly_name:Contig38"
misc_feature 129142..139215
              /note="assembly_name:Contig39"
misc_feature 139316..151356
              /note="assembly_name:Contig40"
misc_feature 151457..163401
              /note="assembly_name:Contig41"
ORIGIN
Query Match      28.5%; Score 95.6; DB 2; Length 163401;
Best Local Similarity 63.5%; Pred. No. 6.7e-08;
Matches 146; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 106 TATGACTATTTTACGTCATATACAAATTTTATGTTTGTAAATGACTTTATGTTGT 165
Db 154116 TTTCCTTTTCTCTCTATTTTATACATCTTTATTTATTTTATTTTACCTTTTGTCTC 154175
QY 166 CATATGATATTTTATGTCATAGAACAAATTTTATGCTTGTGATATGACTTTATGTTA 225
Db 154176 AATCACATATATATATACATATATATATATATATATATATATATATATATATATA 154235
QY 226 TATGGCTATACAACTAGATTTTTTTTGTGTTTGTGACCGAGCTTACTCTGTCACCCAGG 285
Db 154236 TATACATATATATATATATATTTTTTTTTTTTGTGACAGAGTCTCACTCTGTCACCCAGG 154295
QY 286 CTGAGTGTAAATGGCATGGTCTCAGCTCACTGCAACCTCCGCGCTCCCGGG 335
Db 154296 CTGAAGTGCAGTGGCAAAATCTGGCTCGCTGCAACTTCCACCTCCCGAGG 154345
Search completed: May 21, 2004, 20:49:23
Job time : 1350.31 secs

```

